```
TITLE       The three-dimensional solution structure of NaD1, a new floral
            defensin from Nicotiana alata and its application to a homology
            model of the crop defense protein alfAFP
JOURNAL     J. Mol. Biol. 325 (1), 175-188 (2003)
MEDLINE     22361678
PUBMED      12473460
REFERENCE   2 (bases 1 to 601)
AUTHORS     Lay,F.T., Brugliera,F. and Anderson,M.A.
TITLE       Isolation and Properties of Floral Defensins from Ornamental
            Tobacco and Petunia
JOURNAL     Plant Physiol. 131 (3), 1283-1293 (2003)
PUBMED      12644678
REFERENCE   3 (bases 1 to 601)
AUTHORS     Lay,F.T. and Anderson,M.A.
TITLE       Direct Submission
JOURNAL     Submitted (06-MAY-2002) Biochemistry, La Trobe University,
            Bundoora, VIC 3086, Australia
FEATURES             Location/Qualifiers
     source          1..601
                     /organism="Nicotiana alata"
                     /mol_type="mRNA"
                     /db_xref="taxon:4087"
     gene            1..601
                     /gene="NaD1"
     CDS             61..378
                     /gene="NaD1"
                     /note="NaD1, gamma-thionin; possesses antifungal activity;
                     targeted to the vacuole"
                     /codon_start=1
                     /product="flower-specific defensin precursor"
                     /protein_id="AAN70999.1"
                     /db_xref="GI:25005097"
                     /translation="MARSLCFMAFAILAMMLFVAYEVQARECKTESNTFPGICITKPP
                     CRKACISEKFTDGHCSKILRRCLCTKPCVFDEKMKTGAEILAEEAKTLAAALLEEEI
                     MDN"
     mat_peptide     136..276
                     /gene="NaD1"
                     /product="flower-specific defensin"
BASE COUNT      192 a    95 c   115 g   199 t
ORIGIN

Query Match     99.7%;  Score 539.4;  DB 8;  Length 601;
Best Local Similarity   99.8%;  Pred. No. 6.3e-102;
Matches  540;  Conservative    0;  Mismatches   1;  Indels    0;  Gaps    0;

Qy          1 ATGGCTCGCTCCTTGTGCTTCATGGCAGTTGTATCTGTAGGCAAGGATGCTCTTTGTTGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGCTCGCTCCTTGTGCTTCATGGCAGTTGTATCTGTAGGCAAGGATGCTCTTTGTTGCC 120

Qy         61 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAGAAGCAACACATTTCCTGGAATATGCATT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAGAAGCAACACATTTCCTGGAATATGCATT 180

Qy        121 ACCAAAACCACCACAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCAAAACCACCACAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 240

Qy        181 AAAATCCTTCGAGAAGGTGCCTATGTACTACTAAGCCATGTGTGTTGATGAAGATGACTAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAATCCTTCGAGAAGGTGCCTATGTACTACTAAGCCATGTGTGTTGATGAAGATGACTAA 300

Qy        241 ACAGGAGCTGAAAATTTGGCTGAGGAAGCAAAAACACTTGGCTGCTGCAGCTTTGCTTGAAGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAGGAGCTGAAAATTTGGCTGAGGAAGCAAAAACACTTGGCTGCTGCAGCTTTGCTTGAAGAA 360

Qy        301 GAGATAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGATGCAGTATCACACATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGATAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGATGCAGTATCACACATAAT 420

Qy        361 AAAGTTTCTACCTTTCTTTAAAAGTGTAGCTAAGTCTTTTAGTTAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGTTTCTACCTTTCTTTAAAAGTGTAGCTAAGTCTTTTAGTTAGC 480

Qy        421 AAAAGTTTCTACCTTTCTTTAAAAGTGTTTTTAAATTGGCTTTTAGTTAGC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTTTATTACACTTTAAATAAGTGTGGCCACTTCAATCCTTCAATCCTTGCACTAAGT 540
```

RESULT 2

```
LOCUS       SFSTMRNA              589 bp    mRNA    linear
DEFINITION  Solanaceae mRNA for flower-specific thionin.
ACCESSION   Z11748 S41631
VERSION     Z11748.1  GI:21212
KEYWORDS    thionin.
SOURCE      Solanaceae
  ORGANISM  Solanaceae
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales.
REFERENCE   1 (bases 1 to 589)
AUTHORS     Gu,Q., Kawata,E.E., Morse,M.J., Wu,H.M. and Cheung,A.Y.
TITLE       A flower-specific cDNA encoding a novel thionin in tobacco
JOURNAL     Mol. Gen. Genet. 234 (1), 89-96 (1992)
MEDLINE     92357021
PUBMED      1495489
REFERENCE   2 (bases 1 to 589)
AUTHORS     Gu,Q.
TITLE       Direct Submission
JOURNAL     Submitted (25-FEB-1992) Qing GU, Department of Biology, Yale
            University, 165 Prospect St, New Haven, CT, 06511, USA
FEATURES             Location/Qualifiers
     source          1..589
                     /organism="Solanaceae"
                     /mol_type="mRNA"
                     /db_xref="taxon:4070"
                     /clone="49A"
                     /tissue_type="flower of tobacco"
                     /dev_stage="flowering"
     CDS             49..366
                     /codon_start=1
                     /product="flower-specific thionin"
                     /protein_id="CAA77806.1"
                     /db_xref="GI:21213"
                     /db_xref="SWISS-PROT:P32026"
                     /translation="MARSLCFMAFAILAMMLFVAYEVQARECKTESNTFPGICITKPP
                     CRKACISEKFTDGHCSKLLRRCLCTKPCVFDEKMIKTGAEILVEEAKTLAAALLEEEI
                     MDN"
BASE COUNT      184 a    97 c   114 g   194 t
ORIGIN

Query Match     93.2%;  Score 504.2;  DB 8;  Length 589;
Best Local Similarity   95.7%;  Pred. No. 1.3e-94;
Matches  518;  Conservative    0;  Mismatches  23;  Indels    0;  Gaps    0;

Qy          1 ATGGCTCGCTCCTTGTGCTTCATGGCAGTTGTATCTGTAGGCAAGGATGCTCTTTGTTGCC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGGCTCGCTCCTTGTGCTTCATGGCAGTTGTATCTGTAGGCAAGGATGCTCTTTGTTGCC 108

Qy         61 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAGAAGCAACACATTTCCTGGAATATGCATT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAGAAGCAATACAATTCCCTGGAATATGCATT 168

Qy        121 ACCAAAACCACCATGCAGAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 180
              |||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||
Db        169 ACCAAAACCACCATGCAGAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 228

Qy        181 AAAATCCTTCGAGAAGGTGCCTATGTACTACTAAGCCATGTGTGTTGATGAGAAGATGACTAAA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3

```
LOCUS       AB005266                 558 bp    mRNA      linear    PLN 05-FEB-1999
DEFINITION  Nicotiana excelsior mRNA for gamma-thionin, complete cds.
ACCESSION   AB005266
VERSION     AB005266.1  GI:2244704
KEYWORDS    gamma-thionin.
SOURCE      Nicotiana excelsior
  ORGANISM  Nicotiana excelsior
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (sites)
  AUTHORS   Yamada,S.; Komori,T. and Imaseki,H.
  TITLE     cDNA cloning of gamma-thionin from Nicotiana excelsior (Accession
            No. AB005266) (PGR97-131)
  JOURNAL   Plant Physiol. 115, 314 (1997)
REFERENCE   2  (bases 1 to 558)
  AUTHORS   Yamada,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-1997) Shigehiro Yamada, Plant Breeding and
            Genetics Research Laboratory, Japan Tobacco Inc.; 700 Higashibara,
            Iwata, Shizuoka 438, Japan
            (E-mail:Shigehiro.Yamada@pbgrl.jti.co.jp, Tel:0538-32-8700)
            Fax:0538-32-8700)
FEATURES             Location/Qualifiers
     source          1..558
                     /organism="Nicotiana excelsior"
                     /mol_type="mRNA"
                     /db_xref="taxon:61185"
     gene            1..558
                     /gene="NeThio2"
     CDS             33..350
                     /gene="NeThio2"
                     /codon_start=1
                     /product="gamma-thionin"
                     /protein_id="BAA21114.1"
                     /db_xref="GI:2244705"
                     /translation="MARSVCFMAFAILAVMLFVAYDVEAKDCKIESNTPGICITKPP
                     CRKACIKEKFTDGHCSKILRRCLCTKPCVFDEKMIKTGAETLAERATTLAAALIEEI
                     MDN"
```

Qy   229 AAAACTCCTCAGACGGTCGCTATGCACTAAGCCATGTGTTGTTTGATGAGAAGAATGATAAAA 288
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 AAAACTCCTCAGACGGTCGCTATGCACTAAGCCATGTGTTGTTTGATGAGAAGAATGATAAAA 288

Qy   241 ACAGGAGGCTGAAAATTTTGGCTGAGGAGGAAGCAAAAACTTTGGCTGTGCAGCTTTGCCTTGAAGAA 300
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 ACAGGAGGCTGAAACTTTGGTTGAGGAAGCAAAAACTTTGGCTGTGCAGCTTTGCCTTGAAGAA 348

Qy   301 GAGATAATGGATAACTAATTAGAGAGAGAAAATTAAGGATGACAGTATCACATAAT 360
          |||||||||||| |||||||||||||||||||||||||||||| | | |||| |
Db   349 GAGATAATGGATAACTAATTAGAGAGAGAAAATTAAGGATGACAGTGTCACACATAAT 408

Qy   361 AAAGTTTCTACCTTTCTTAAAAGTGTGTTTTAATTGTTGGCTTTTAGTAGC 420
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   409 AAAGTTGCTACCTTTCTTAAAAGTGTGTTTTAATTGTTGGCTTTTAGTAGC 468

Qy   421 CTTTTATTTACACTTTAAATAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   469 CTTTTGTTTACACTTTAAATAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGT 528

Qy   481 TTATTTGTTGCTACTTTTAATGAAAAATGAAAATGACCTTCTAATGGTCTTTGGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   529 TTATTCGTGTACTTTTAATGAAAAATCACCTTCTATGGTTTTGGTT 588

Qy   541 A 541
          |
Db   589 A 589

RESULT 4

```
Best Local Similarity  94.9%;
Matches 499;  Conservative  0;  Pred. No. 3.5e-90;
                                 Mismatches  27;  Indels  0;  Gaps  0;
```

```
LOCUS       AB005250                 549 bp    mRNA      linear    PLN 19-JUN-1999
DEFINITION  Nicotiana paniculata mRNA for gamma-thionin, complete cds.
ACCESSION   AB005250
VERSION     AB005250.1  GI:2251080
KEYWORDS    gamma-thionin.
SOURCE      Nicotiana paniculata
  ORGANISM  Nicotiana paniculata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (sites)
  AUTHORS   Komori,T.; Yamada,S. and Imaseki,H.
  TITLE     A cDNA clone for gamma-thionin from Nicotiana paniculata (Accession
            No. AB005250) (PGR97-132)
  JOURNAL   Plant Physiol. 115, 314 (1997)
REFERENCE   2  (bases 1 to 549)
  AUTHORS   Komori,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-1997) Toshiyuki Komori, Plant Breeding and
            Genetics Research Laboratory, Japan Tobacco Inc.; 700 Higashibara,
            Iwata, Shizuoka 438, Japan
            (E-mail:Toshiyuki.Komori@pbgrl.jti.co.jp, Tel:0538-32-7116,
            Fax:0538-32-8700)
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Nicotiana paniculata"
                     /mol_type="mRNA"
                     /db_xref="taxon:62141"
     gene            1..549
                     /gene="NpThio1"
     CDS             48..368
```

Qy   1 ATGGGCTCGGCTCCTTGTGCTTCATGGCAATTTGCTATCTTGGCAAGGATGGCTCTTTGTTGTTGCC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGGCTCGGCTCCTTGTGCTTCATGGCAATTTGCTATCTTGGCAAGGATGGCTCTTTGTTGTTGCC   60

Qy   61 TATGAGGTGCAAGCTAGAGAGAATGCAAAACAGAAAAGCAAACAACATTTCCTGGAAATATGGCATT 120
          ||| || ||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   33 ATGGCTCGGCTCCGTGTGCTTCATGGCAATTTGCTATCTTGGCAGTGATGCTCTTTGTTGCC   92

Qy   93 TATGAGTGGAAGCTAAAGATTGCAAAACAGAAAAGCAAACAACAGAAAAGCAAACAACATTTCCTGGAAATATGGCATT 152
          || ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   93 TATGATGTGGAAGCTAAAGATTGCAAAACAGAAAAGCAAACAACAGAAAAGCAAACAACATTTCCTGGAAATATGGCATT 152

Qy   121 ACCAAACCACCANTGCAGAAAAAGCTTGTATCAGTGAGAAATTACTGATGGTCATTGTAGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   153 ACCAAACCACCANTGCAGAAAAAGCTTGTATCAGTGAGAAATTACTGATGGTCATTGTAGC 212

Qy   181 AAAAATCCTCAGAAAGGTCGGTCTATGCTACTAACGCCATGTGTTGTTTGATGAGAAGAAGTGACTAAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   213 AAAAATCCTCAGAAAGGTCGGTCTATGCTACTAACGCCATGTGTTGTTTGATGAGAAGAAGTGACTAAA 272

Qy   241 ACAGGAGGCTGAAAAATTTTGGCTGCGAGGAAGCAAAAACTTTGGCTGCGCAGCTTTGCCTTGAAGAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   273 ACAGGAGGCTGAAAACTTTAGCTGCGAGGAAGCAACAACTTTGGCTGCGCAGCTTTGCTTGAAGAA 332

Qy   301 GAGATAATGGATAACTAATTAGAGATTAGAGAGAAAATTAAGGATGCAGTAATCACACATAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   333 GAGATAATGGATAACTAATTAGAGATTAGAAATTAAGGATGAGAGTCACACATAAT 392

Qy   361 AAAGTTTCTACCTTTCTTAAAAAGTGTAGGCTAAATGTGTGTTTAATTGGCTTTTAGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   393 AAAGTTTCTACCTTTCTTAAAAAGTGTAGGCTAAATGTGTGTTTAATTGGCTTTTAGTAGC 452

Qy   421 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   453 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGT 512

Qy   481 TTATTTGTGTACTTTTAATGAAAAATGACCTTCTATGGTTCTTGGTT 526
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   513 TTATTCGTGTACTTTTAATGAAAAATCACCTTCTATGGTTTTGTTT 558

```
BASE COUNT    170 a    94 c   108 g   186 t
ORIGIN
Query Match   89.2%;   Score 482.8;  DB 8;  Length 558;
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6

Run on: January 28, 2004, 08:46:51 ; Search time 14 Seconds
                                           (without alignments)
                                           352.700 Million cell updates/sec

Title: US-10-072-809B-18
Perfect score: 552
Sequence: 1 MARSLCFMAFAILARMLFVA...........LAEEAKTLAAALLEEEIMDN 105

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	529	95.8	105	1	THG1_TOBAC	P32026 nicotiana t
2	468.5	84.9	106	1	THG1_NICPA	O24115 nicotiana p
3	144	26.1	106	1	THG_PETIN	Q40901 petunia int
4	143.5	26.0	78	1	THGI_PETIN	Q9zu18 arabidopsis
5	138.5	25.1	74	1	P322_SOLTU	P20346 solanum tub
6	137.5	24.9	77	1	THG1_ARATH	Q39182 arabidopsis
7	137.5	24.9	77	1	THG3_ARATH	Q9zu17 arabidopsis
8	137	24.8	78	1	THG_HELAN	P82659 helianthus
9	135	24.5	74	1	DEF2_CAPAN	O65740 capsicum an
10	132	23.9	52	1	DEF2_SPIOL	P81571 spinacia ol
11	127.5	23.1	77	1	THG2_ARATH	Q41914 arabidopsis
12	123	22.3	75	1	DEF1_CAPAN	Q43413 capsicum an
13	112	20.3	80	1	AFP2_ARATH	O80995 arabidopsis
14	111	20.1	74	1	SRP_SOYBN	Q07502 glycine max
15	110	19.9	80	1	AFP1_BRANA	P30225 brassica na
16	108	19.6	80	1	AFP4_RAPSA	O24331 raphanus sa
17	107	19.4	47	1	THG1_MAIZE	P81008 zea mays (m
18	106.5	19.4	79	1	AFP3_RAPSA	O24332 raphanus sa
19	105	19.1	79	1	AFP3_ARATH	O80994 arabidopsis
20	105	19.0	80	1	AFP3_ARATH	O09198 arabidopsis
21	103	18.7	47	1	SA21_SORBI	P21925 sorghum bic
22	103	18.7	80	1	SIA3_SORBI	P30224 arabidopsis
23	103	18.7	80	1	AFP1_ARATH	P30230 raphanus sa
24	102.5	18.6	79	1	AFP2_RAPSA	Q39313 brassica na
25	102	18.6	47	1	AFP3_BRANA	P83399 vigna ungui
26	101	18.3	47	1	THGC_VIGUN	P81456 vicia faba
27	101	18.3	47	1	FAB1_VICFA	P81457 vicia faba
28	99	17.9	47	1	FAB2_VICFA	P20230 hordeum vul
29	99	17.9	48	1	THG_HORVU	P21924 sorghum bic
30	97	17.6	80	1	SIA2_SORBI	Q9fi23 arabidopsis
31	94.5	17.1	74	1	AFP4_ARATH	Q01784 pisum sativ
32	94	17.0	47	1	DR39_PEA	Q20159 triticum ae
33	92.5	16.8	75	1	THG2_WHEAT	P18646 vigna ungui
34	88.5	16.0	51	1	AF2A_SINAL	P30232 sinapis alb
35	86.5	15.9	47	1	THG1_WHEAT	P20158 triticum ae
36	86.5	15.7	72	1	D230_PEA	Q01783 pisum sativ
37	85.5	15.5	46	1	AX2_BETVU	P82010 beta vulgar
38	85.5	15.5	46	1	AFP1_SINAL	P30231 sinapis alb
39	84.5	15.3	46	1	AX1_BETVU	P81493 beta vulgar
40	78	14.1	47	1	SIA1_SORBI	P21923 sorghum bic
41	77	13.9	70	1	DMYC_DROME	P41964 drosophila
42	76	13.8	47	1	PSD2_PEA	P81930 pisum sativ
43	76	13.8	47	1	TH22_MAIZE	P81009 zea mays (m
44	73.5	13.3	297	1	SG54_DROME	O00725 drosophila
45	68	12.3	322	1	FSA_BRARE	Q9yhv4 brachydanio

ALIGNMENTS

RESULT 1
THGF_TOBAC

ID   THGF_TOBAC   STANDARD;   PRT;   105 AA.
AC   P32026;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flower-specific gamma-thionin precursor.
GN   FST.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. W38; TISSUE=Flower buds;
RX   MEDLINE=92357021; PubMed=1495489;
RA   Gu Q., Kawata E.E., Morse M.-J., Wu H.-M., Cheung A.Y.;
RT   "A flower-specific cDNA encoding a novel thionin in tobacco.";
RL   Mol. Gen. Genet. 234:89-96(1992).
CC   -!- FUNCTION: INVOLVED IN FLORAL ORGANOGENESIS. MAY PLAY A PROTECTIVE
CC   ROLE-IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM
CC   POTENTIAL PATHOGEN ATTACK.
CC   -!- SUBCELLULAR LOCATION: POSSIBLY THE CELL WALL OR VACUOLE.
CC   -!- TISSUE SPECIFICITY: FLOWER. FOUND IN PETALS, STAMEN AND PISTILS,
CC   BUT NOT IN SEPALS. IN PARTICULAR, ACCUMULATION IN A CONFIGURATION
CC   SURROUNDING THE INNER REPRODUCTIVE WHORLS.
CC   -!- DEVELOPMENTAL STAGE: ACCUMULATES IN DEVELOPING FLOWERS AND ITS
CC   LEVEL DROPS AS FLOWERS MATURE.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; Z11748; CAA77806.1; -.
DR   PIR; S23574; S23574.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Plant defense; Plant toxin; Signal.
FT   SIGNAL   1   25   POTENTIAL.
FT   CHAIN   26   105   FLOWER-SPECIFIC GAMMA-THIONIN.
FT   DISULFID   28   72   BY SIMILARITY.
FT   DISULFID   39   59   BY SIMILARITY.
FT   DISULFID   45   66   BY SIMILARITY.
FT   DISULFID   49   68   BY SIMILARITY.
SQ   SEQUENCE   105 AA;   11750 MW;   DA7F451F6651BC3 CRC64

```
Qy         Query Match
           Best Local Similarity   95.8%;   Score 529;  DB 1;  Length 105;
           Conservative  95.2%;  Pred. No. 2.6e-49;
Matches 100;  Conservative   1;  Mismatches  4;  Indels  0;  Gaps  0;

Qy     1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCPCRKACISEKFTDGHCS  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MARSLCFMAFAILAAVMLFVAYEVQARECKIESNTFPGICITKPPCRKACISEKFTDGHCS  60

Qy    61  KILRRCLCTKPCVFDEKMTKTGAEIIAEEAKTLAAALLEEIMDN  105
          |||||||||||||||||||||||||||||||||||||||||||||
Db    61  KILRRCLCTKPCVFDEKMIKTGAETLVEEAKTLAAALLEEIMDN  105
```

RESULT 2

```
ID     THG1_NICPA      STANDARD;       PRT;     106 AA.
AC     O24115;
DT     15-DEC-1998 (Rel. 37, Created)
DT     15-DEC-1998 (Rel. 37, Last sequence update)
DT     28-FEB-2003 (Rel. 41, Last annotation update)
DE     Gamma-thionin 1 precursor.
GN     THIO1.
OS     Nicotiana paniculata.
OC     Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC     Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX     NCBI_TaxID=62141;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     TISSUE=Leaf;
RA     Komori T., Yamada S., Imaseki H.;
RT     "A cDNA clone for gamma-thionin from Nicotiana paniculata.";
RL     (In) Plant Gene Register PGR97-132.
CC     -!- INDUCTION: By salt stress.
CC     -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC     --------------------------------------------------------
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC     or send an email to license@isb-sib.ch).
CC     --------------------------------------------------------
DR     EMBL; AB005250; BAA21325.1; -.
DR     InterPro; IPR002118; Gamma-thionin.
DR     InterPro; IPR003614; Knot1.
DR     Pfam; PF00304; Gamma-thionin; 1.
DR     ProDom; PD002594; G_Purothionin; 1.
DR     SMART; SM00505; Knot1; 1.
DR     PROSITE; PS00940; GAMMA_THIONIN; FALSE_NEG.
KW     Plant defense; Plant toxin; Signal.
FT     SIGNAL      1     25        POTENTIAL.
FT     CHAIN      26    106        GAMMA-THIONIN 1.
FT     DISULFID   29     73        BY SIMILARITY.
FT     DISULFID   40     60        BY SIMILARITY.
FT     DISULFID   46     67        BY SIMILARITY.
FT     DISULFID   50     69        BY SIMILARITY.
SQ     SEQUENCE   106 AA;  11722 MW;  2698S0E11GD7775A4 CRC64;
```

```
Qy         Query Match
           Best Local Similarity   84.9%;   Score 468.5;  DB 1;  Length 106;
           Conservative   84.0%;  Pred. No. 6.7e-43;
Matches 89;  Conservative   8;  Mismatches  8;  Indels  1;  Gaps  1;

Qy     1  MARSLCFMAFAILARMLFVAYEVQOARE-CKTESNTFPGICITKPPCPCRKACISEKFTDGHC  59
          ||||||||||||| |||||||| |||| ||||||||||||||||||||||||||||||||
Db     1  MARSLCFMAFAFVLAAMLFVAYEVQAKSTCKAESNTFPGLCITKPPCPCRKACISEKFTDGKC  60

Qy    60  SKILRRCL.CTKPCVFDEKMTKTGAEIIAEEAKTLAAALLEEIMDN  105
          ||||||:|  ||||||||||| |||:|| ||||||||||||||:|||
Db    61  SKILRRCICYKPCVFDGKMIQTGAENLAEBAEFILAAALLEEENMDN  106
```

RESULT 3

```
ID     THG_PETIN       STANDARD;       PRT;      78 AA.
AC     Q40501;
DT     15-DEC-1998 (Rel. 37, Created)
DT     15-DEC-1998 (Rel. 37, Last sequence update)
DT     28-FEB-2003 (Rel. 41, Last annotation update)
DE     Gamma-thionin homolog PPT precursor.
OS     Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).
OC     Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC     Asteridae; lamiids; Solanales; Solanaceae; Petunia.
OX     NCBI_TaxID=4103;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     TISSUE=Pistil;
RX     MEDLINE=95036017; PubMed=7948892;
RA     Karunanandaa B., Singh A., Kao T.H.;
RT     "Characterization of a predominantly pistil-expressed gene encoding a
RT     gamma-thionin-like protein of Petunia inflata.";
RL     Plant Mol. Biol. 26:459-464(1994).
CC     -!- FUNCTION: May be involved in the defense of the pistil against
CC         pathogen infection.
CC     -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE PISTIL DURING
CC         ALL STAGES OF FLOWER DEVELOPMENT.
CC     -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC     --------------------------------------------------------
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC     or send an email to license@isb-sib.ch).
CC     --------------------------------------------------------
DR     EMBL; L27173; AAA64740.1; -.
DR     PIR; S52634; S52634.
DR     HSSP; P20230; 1GPT.
DR     InterPro; IPR002118; Gamma-thionin.
DR     InterPro; IPR003614; Knot1.
DR     Pfam; PF00304; Gamma-thionin; 1.
DR     ProDom; PD002594; G_Purothionin; 1.
DR     SMART; SM00505; Knot1; 1.
DR     PROSITE; PS00940; GAMMA_THIONIN; 1.
KW     Plant defense; Signal.
FT     SIGNAL      1     31        POTENTIAL.
FT     CHAIN      32     78        GAMMA-THIONIN HOMOLOG PPT.
FT     DISULFID   34     78        BY SIMILARITY.
FT     DISULFID   45     65        BY SIMILARITY.
FT     DISULFID   51     72        BY SIMILARITY.
FT     DISULFID   55     74        BY SIMILARITY.
SQ     SEQUENCE    78 AA;  8799 MW;  263DB3AC9D619265 CRC64;
```

```
Qy         Query Match
           Best Local Similarity   26.1%;   Score 144;  DB 1;  Length 78;
           Conservative   35.9%;  Pred. No. 1.2e-08;
Matches 28;  Conservative  14;  Mismatches  30;  Indels  6;  Gaps  1;

Qy     1  MARSLCFMAFAILARMLFVAYEV----------QARECKTESNTFPGICITKPPCPCRKACISEKF  54
          ||||| |||:|| ||||||||||          ||:||:||:||||||:|:|||||||||||:|
Db     1  MGRSIRLFAIFFLIAMLFLSTEMGPMTSAEARTCESQSHRFHGTCVRESNCASVCQTEGF  60

Qy    55  TDGHCSKILRRCLCTKPC  72
          :|:| ||:|||||:|:|
Db    61  IGGNCRAFRRCPCTRNC  78
```

RESULT 4

```
ID     THG4_ARATH      STANDARD;       PRT;      73 AA.
AC     Q9ZUL8;
DT     16-OCT-2001 (Rel. 40, Created)
DT     16-OCT-2001 (Rel. 40, Last sequence update)
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 09:22:12 ; Search time 1169 Seconds
                 (without alignments)
                 1655.150 Million cell updates/sec

Title: US-10-072-809B-17
Perfect score: 541
Sequence: 1 atggctcgctccttgtgctt..........tggttaaaaaaaaaaaaa 541

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 2356869 seqs, 1788225258 residues

Total number of hits satisfying chosen parameters: 4713738

Post-processing: Minimum Match 0%
                 Maximum Match 100%
Minimum DB seq length: 0
Maximum DB seq length: 200000000
                 Listing first 45 summaries

Database :   Published Applications_NA:*
          1: /cgn2_5/ptcdata/1/pubpna/US07_PUBCOMB.seq:*
          2: /cgn2_6/ptcdata/1/pubpna/PCT_NEW_PUB.seq:*
          3: /cgn2_6/ptcdata/1/pubpna/US06_NEW_PUB.seq:*
          4: /cgn2_6/ptcdata/1/pubpna/US06_PUBCOMB.seq:*
          5: /cgn2_6/ptcdata/1/pubpna/US07_NEW_PUB.seq:*
          6: /cgn2_6/ptcdata/1/pubpna/PCTUS_PUBCOMB.seq:*
          7: /cgn2_6/ptcdata/1/pubpna/US08_NEW_PUB.seq:*
          8: /cgn2_6/ptcdata/1/pubpna/US09A_PUBCOMB.seq:*
          9: /cgn2_6/ptcdata/1/pubpna/US09B_PUBCOMB.seq:*
         10: /cgn2_6/ptcdata/1/pubpna/US09C_PUBCOMB.seq:*
         11: /cgn2_6/ptcdata/1/pubpna/US09_NEW_PUB.seq:*
         12: /cgn2_6/ptcdata/1/pubpna/US09_NEW_PUB.seq2:*
         13: /cgn2_6/ptcdata/1/pubpna/US10A_PUBCOMB.seq:*
         14: /cgn2_6/ptcdata/1/pubpna/US10B_PUBCOMB.seq:*
         15: /cgn2_6/ptcdata/1/pubpna/US60_NEW_PUB.seq:*
         16: /cgn2_6/ptcdata/1/pubpna/US60_NEW_PUB.seq:*
         17: /cgn2_6/ptcdata/1/pubpna/US60_PUBCOMB.seq:*
         18: /cgn2_6/ptcdata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	DB ID	Description
1	541	100.0	541	13	US-10-072-809A-17	Sequence 17, Appl
2	240	44.4	240	13	US-10-072-809A-15	Sequence 15, Appl
3	223	41.2	223	13	US-10-072-809A-19	Sequence 19, Appl
4	216	39.9	216	13	US-10-072-809A-13	Sequence 13, Appl
5	189.8	35.1	564	9	US-09-777-347-1	Sequence 1, Appl
6	147.4	27.2	3528	9	US-09-777-347-2	Sequence 2, Appli
7	141	26.1	141	13	US-10-072-809A-7	Sequence 7, Appl
8	99	18.3	99	13	US-10-072-809A-11	Sequence 11, Appl
9	75	13.9	75	13	US-10-072-809A-9	Sequence 9, Appl
10	48	8.9	3673778	13	US-10-312-841-1	Sequence 1, Appli
11	46.2	8.5	516	15	US-10-178-213-82	Sequence 82, Appl
12	44.6	8.2	508	15	US-10-178-213-85	Sequence 85, Appl
13	44.6	8.2	578	15	US-10-178-213-292	Sequence 292, App
14	44.6	8.2	13326	13	US-10-311-455-1685	Sequence 1685, Ap
15	43.4	8.0	283	9	US-09-294-093B-4663	Sequence 4663, Ap
16	42.4	7.8	6042	13	US-10-311-455-1565	Sequence 1565, Ap
17	42.2	7.8	297	9	US-09-294-093B-4741	Sequence 4741, Ap
18	42.2	7.8	559	15	US-10-178-213-1	Sequence 1, Appli
19	42	7.8	459	15	US-10-178-213-322	Sequence 322, App
20	42	7.8	1692	13	US-10-349-680-155	Sequence 155, App
21	42	7.8	822900	10	US-09-790-988-1	Sequence 1, Appli
22	41.8	7.7	640681	10	US-10-178-21-4	Sequence 4, Appli
23	41.8	7.7	6137	13	US-10-240-485-4	Sequence 4, Appli
24	41.6	7.7	1438	10	US-09-938-842A-4595	Sequence 4595, Ap
25	41.6	7.7	1438	15	US-09-938-842A-4595	Sequence 4595, Ap
26	41.2	7.6	468	15	US-10-178-213-316	Sequence 316, App
27	41.2	7.6	640681	12	US-10-239-676-3	Sequence 3, Appli
28	41	7.6	640	13	US-10-311-455-351	Sequence 351, App
29	41	7.6	6665	13	US-10-292-798-1393	Sequence 1393, App
30	41	7.6	6665	15	US-09-814-355-55	Sequence 55, Appli
31	41	7.6	6665	13	US-10-311-455-5	Sequence 5, Appli
32	41	7.6	10433	13	US-10-311-455-351	Sequence 351, App
33	40.8	7.5	608	13	US-09-814-355-5190	Sequence 5190, Ap
34	40.8	7.5	608	13	US-10-311-455-1809	Sequence 1809, Ap
35	40.8	7.5	6089	13	US-10-311-455-11478	Sequence 11478, A
36	40.6	7.5	19659	13	US-10-311-455-739	Sequence 739, Appli
37	40.6	7.5	335913	11	US-09-754-853A-2	Sequence 2, Appli
38	40.6	7.5	335913	11	US-09-754-853A-3	Sequence 3, Appli
39	40.6	7.5	3673778	13	US-09-312-841-1	Sequence 1, Appli
40	40.4	7.5	539	13	US-09-814-353-17862	Sequence 17862, A
41	40.2	7.4	6074	13	US-10-240-453-248	Sequence 248, App
42	40.2	7.4	6131	13	US-10-311-455-863	Sequence 863, App
43	40	7.4	4190	13	US-10-311-455-415	Sequence 415, App
44	39.8	7.4	550	15	US-10-178-213-424	Sequence 424, App
45	39.8	7.4	9219	13	US-10-311-455-2399	Sequence 2399, Ap

ALIGNMENTS

RESULT 1
US-10-072-809A-17
Sequence 17, Application US/10072809A

Publication No. US20030217382A1
GENERAL INFORMATION:
   APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
   TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and ...
   TITLE OF INVENTION: therefor
   FILE REFERENCE: 18-01
   CURRENT APPLICATION NUMBER: US/10/072,809A
   CURRENT FILING DATE: 2002-09-12
   PRIOR APPLICATION NUMBER: USSN 60/267,271
   PRIOR FILING DATE: 2001-02-08
   NUMBER OF SEQ ID NOS: 61
   SOFTWARE: PatentIn version 3.0
   SEQ ID NO 17
   LENGTH: 541
   TYPE: DNA
   ORGANISM: Nicotiana alata
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1)..(318)
US-10-072-809A-17

  Query Match   100.0%; Score 541; DB 13; Length 541;
  Best Local Similarity   100.0%; Pred. No. 9.6e-126;
  Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGGCTCGCTCCTTGTGCTTCATGGGCATTTGCTACTCTTGGCAAGGATGGCTCTTTGTTGCC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCTCGCTCCTTGTGCTTCATGGGCATTTGCTACTCTTGGCAAGGATGGCTCTTTGTTGCC 60

Qy 61 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAAGCAACACACATTTCCTGGAATATGCATT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAAGCAACACACATTTCCTGGAATATGCATT 120

Qy 121 ACCAAAACCACCATGCAGAGAAAAAGCTTGTATCACGTGAGAAAATTTACTGATGGTCATTGTGAGC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Left column (RESULT 2)

Db 121 ACCAAACACCATGCAGAAAAGCTTGTATCAGTGAGGAAAATTACTCAGTGACTGTAGC 180

Qy 181 AAAATCCTCAGAAAGGTGCCTATGTACTACTAAGCCATGTGTTGATGAGATGACTAAA 240
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AAAATCCTCAGAAAGGTGCCTATGTACTACTAAGCCATGTGTTGATGAGATGACTAAA 240

Qy 241 ACAGGAGCTGAAATTTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAAGAA 300
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ACAGGAGCTGAAATTTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAAGAA 300

Qy 301 GAGATAATGGGATAACTAATTACAGATTAGAAGAAGAAATTAAGGATGACTCAAA 360
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GAGATAATGGGATAACTAATTACAGATTAGAAGAAGAAATTAAGGATGACTCAAA 360

Qy 361 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTTGCTTTAGTTAGCTAGC 420
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTTGCTTTAGTTAGCTAGC 420

Qy 421 CTTTTATTTACACTTTAAATAAGTGTGGCCACTTCAATCCTTTGTGCAATCTTGCACTAAGT 480
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CTTTTATTTACACTTTAAATAAGTGTGGCCACTTCAATCCTTTGTGCAATCTTGCACTAAGT 480

Qy 481 TTATTTGTGTACTTTTAATGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAA 540
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 TTATTTGTGTACTTTTAATGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAA 540

Qy 541 A 541
   |
Db 541 A 541

RESULT 2

RESULT 2
US-10-072-809A-15
; Sequence 15, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0.
; SEQ ID NO 15
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Nicotiana alata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(240)
US-10-072-809A-15

Query Match 44.4%; Score 240; DB 13; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.2e-50;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 AGAGAATGCAAAACAGAAAAGCAAACACATTCCTGGAATATGCAATTACCAAAACCACCATGC 135
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 AGAGAATGCAAAACAGAAAAGCAAACACATTCCTGGAATATGCAATTACCAAAACCACCATGC 60

Qy 136 AGAAAAGCTTGTATCAGTGAGGAAAATTACTGATGGTCATTGTAGCAAAATCCTCAGGAAGG 195
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 AGAAAAGCTTGTATCAGTGAGGAAAATTACTGATGGTCATTGTAGCAAAATCCTCAGGAAGG 120

Qy 196 TGCCTATGTACTAAGGCCATGTGTTGATGAGAAGAGATGACTAAAAACAGGAGCTGAAATT 255
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TGCCTATGTACTAAGCCATGTGTTGATGAGAAGAGATGACTAAAAACAGGAGCTGAAATT 180

Qy 256 TTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAAGAAGAGATAATGGGATAAC 315
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 TTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAAGAAGAGATAATGGGATAAC 240

RESULT 3

RESULT 3
US-10-072-809A-19
; Sequence 19, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 223
; TYPE: DNA
; ORGANISM: Nicotiana alata
US-10-072-809A-19

Query Match 41.2%; Score 223; DB 13; Length 223;
Best Local Similarity 100.0%; Pred. No. 3.9e-46;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 319 TTAGACATTAGAAGAAGAAATTAAGGATGACAGCAGTATCACACATAATAAAGTTTCACTTTAAA 378
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 TTAGACATTAGAAGAAGAAATTAAGGATGACAGCAGTATCACACATAATAAAGTTTCACTTTAAA 60

Qy 379 AAAAGTTGTAGCTAATGTTGTGTTTGTTTTTATGGCCTTTTAGTTAGCCCTTTATTACACTTTAAA 438
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 AAAAGTTGTAGCTAATGTTGTGTTTGTTTTTATGGCCTTTTAGTTAGCCCTTTATTACACTTTAAA 120

Qy 439 TAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGTTTATTTGTGTACTTTTAA 498
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TAAGTGTGGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGTTTATTTGTGTACTTTTAA 180

Qy 499 TGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAAAAAAAA 541
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 TGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAAAAAAAA 223

RESULT 4

RESULT 4
US-10-072-809A-13
; Sequence 13, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Nicotiana alata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(216)
US-10-072-809A-13

Query Match 39.9%; Score 216; DB 13; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.2e-44;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 TTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAAGAAGAGATAATGGGATAAC 240

```
Qy      1 ATGGCTCGGCTCCTTGTGCTTCAICAGGCATTGCTATCTTGGCAAGGATGCTCTCTTGTTGCC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTCGGCTCCTTGTGCTTCAICAGGCATTGCTATCTTGGCAAGGATGCTCTCTTGTTGCC    60

Qy     61 TATGAGGTGCAAGCTAGAGAATCCAAAACACGAAAGCAACACATTCCTGGAATATGCCATT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TATGAGGTGCAAGCTAGAGAATCCAAAACACGAAAGCAACACATTCCTGGAATATGCCATT   120

Qy    121 ACCAAACACACCATGCCAGAAAAAGCTGTATTCAGTGAGAAATTAACTGATGGTCATTGTTAGC   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCAAACACACCATGCCAGAAAAAGCTTGTATTCAGTGAGAAATTTACTGATGGTCATTGTTAGC   180

Qy    181 AAAAATCCTCAGAAGGTGCCTAIGTTACTAAGCCCATGT   216
          |||||||||||||||||||||||||||||||||||||||
Db    181 AAAAATCCTCAGAAGGTGCCTAIGTTACTAAGCCCATGT   216
```

RESULT 5
US-09-777-347-1
; Sequence 1, Application US/09777347
; Patent No. US20010014977A1
GENERAL INFORMATION:
  APPLICANT: McBride, Kevin E.
            Stalker, David M.
  TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Calgene, Inc.
    STREET: 1920 Fifth Street
    CITY: Davis
    STATE: CA
    COUNTRY: USA
    ZIP: 95616
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
    COMPUTER: Apple Macintosh
    OPERATING SYSTEM: Macintosh 7.1
    SOFTWARE: Microsoft Word 5.1 (a)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/777,347
    FILING DATE: 05-Feb-2001
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/487,087
    FILING DATE: <Unknown>
    APPLICATION NUMBER: USSN 07/554,195
    FILING DATE: 17-JUL-90
    APPLICATION NUMBER: USSN 07/382,518
    FILING DATE: 19-JUL-89
  ATTORNEY/AGENT INFORMATION:
    NAME: Carl J. Schwedler
    REGISTRATION NUMBER: 36,924
    REFERENCE/DOCKET NUMBER: CGNE 91-1
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (916) 753-6313
    TELEFAX: (916) 753-1510
  INFORMATION FOR SEQ ID NO: 1
    SEQUENCE CHARACTERISTICS:
      LENGTH: 564 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: cDNA to mRNA
    SEQUENCE DESCRIPTION: SEQ ID NO: 1

US-09-777-347-1

Query Match   35.1%;   Score 189.8; DB 9; Length 564;
Best Local Similarity   68.9%;   Pred. No. 1.3e-37;
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

```
Qy      1 ATGGCTCGTCGTCCTTGTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC    60
          |||||  ||| || ||||| ||||||||||||||||||||| || |||||||  | | ||||||
Db     40 ATGGCTCGTTGCTTCCATTTTCTTCAICATGGCATTTTGGTCTTGGCAATGATGCTCTTCTTGTTTACC    99
```

RESULT 6
US-09-777-347-2
; Sequence 2, Application US/09777347
; Patent No. US20010014977A1
GENERAL INFORMATION:
  APPLICANT: McBride, Kevin E.
            Stalker, David M.
  TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Calgene, Inc.
    STREET: 1920 Fifth Street
    CITY: Davis
    STATE: CA
    COUNTRY: USA
    ZIP: 95616
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
    COMPUTER: Apple Macintosh
    OPERATING SYSTEM: Macintosh 7.1
    SOFTWARE: Microsoft Word 5.1 (a)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/777,347
    FILING DATE: 05-Feb-2001
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/487,087
    FILING DATE: <Unknown>
    APPLICATION NUMBER: USSN 07/554,195
    FILING DATE: 17-JUL-90
    APPLICATION NUMBER: USSN 07/382,518
    FILING DATE: 19-JUL-89
  ATTORNEY/AGENT INFORMATION:
    NAME: Carl J. Schwedler
    REGISTRATION NUMBER: 36,924
    REFERENCE/DOCKET NUMBER: CGNE 91-1
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (916) 753-6313
    TELEFAX: (916) 753-1510
  INFORMATION FOR SEQ ID NO: 2
    SEQUENCE CHARACTERISTICS:
      LENGTH: 3528 base pairs
      TYPE: nucleic acid

```
Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTCCTGGAATATGC   117
          ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db    100 TATGAGGTAGAAGCTCAGCAAAATTTGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGT   159

Qy    118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAGAAATTTACTGATGGTCATTGT   177
          ||||| | |||| |||| ||||||||||||| ||||||||||| |||||||||||||||||||
Db    160 TTTATGGACTCATCATGTAGAAAAATATTGTATCAAAGAGAGAAAATATTGTTGGACATTGT   219

Qy    178 AGCAAAAATCCTCAGAAGGGTGCCTAGTTACTAAAGCCATGTGTGTTTGATGAGAAGATGACT   237
          |||||| |||| |||||||| ||| ||||||||||||| | |||||||||| |||||| ||||
Db    220 AGCAAACTCCAAAGGGAAGTGCTCATGCACTAAAGCCCATGTCATGTTT---TGACAAAATCTCA   276

Qy    238 AAAACAGGAGGCTGAAATTTTGGGCTGAGGAAGCAAAAAACTTTGGCCTGCAGCTTGCTTGAA   297
          ||||| |||||| ||||||||||||||||||||||||| || ||||| ||| |||||||||||
Db    277 AGTGAGTTAAAAGCAACTTTGGGGTGAGGAAGCAAAAAACTCTAAGTGAAGTTGTGCTTGAA   336

Qy    298 GAAGAGGATAATGGGATAACTAATTAGAGATTAGAAGCAAAAATTAAAGGAT-------GCA   346
          ||||| |||| ||||| ||| |||||||||||| || |||||| |||||||||       |||
Db    337 GAAGAGATTATGGATGGATGGTAATTAATTAAGTGAGGTTAAAATAAGGATTTTGAGTTGGTCAAAA   396

Qy    347 GTATCACACATAATAAAAGTTTCTACCTTTCCTTTAAAAAGTGTAGC---TAAATGTTGGTGTTTTT   403
          ||| | ||||| ||||| | | ||| |||| |||| |||||||||||   |||||| || | |||
Db    397 AAAACAAAAATTAATAAAAGTGTTGCCTTTTCTTATTGTGGGTAGCCTTTGGTAGAGTTGTTGTTTAG   456

Qy    404 AATTGGCTTTTAGTAGCCTTTTATTACACTTTAAATAAGTGTGGCCACTTCAAT   456
          ||||| | ||| ||||||| | | |||||||||||| |||| |||||||||||||
Db    457 TATTGGCCTATAGTAGCCATTTGACACCATTAAAATAAGTTTGTTGTGACACATCATT   509
```

```
;
;    STRANDEDNESS: double
;    TOPOLOGY: linear
;    MOLECULE TYPE: genomic DNA
;    SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-777-347-2

Query Match     27.2%;  Score 147.4;  DB 9;  Length 3528;
Best Local Similarity     67.1%;  Pred. No. 1.5e-26;
Matches 261;  Conservative   0;  Mismatches 111;  Indels  17;  Gaps   3;

Qy   82  TGCAAAACAGAAAAGCAACACACATTTCCTGGAATATGCATTACCAAAACACCATGCAGAGAAAA 141
            ||||| |||| || |||  |  |||||| |||||||||| | ||||   | |||| |||||||
Db 2942  TGCAAAGCACCAAGCCAAAACTTTCCCAGGATTATGTTTAATGGACTCATCATGTGAGAAAA 3001

Qy  142  GCTTGTATCAGTGAGAAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
            | ||||||||  |||||||||||||  | | ||||||||||||  ||| |||||  | | |
Db 3002  TATTGTATCAAAGAGAAATTTACTCGGTGAGACATTGTAGCAAAACTCCAAAGAGAAGTGCTCA 3061

Qy  202  TGTACTAAGCCATGTGTTGTTTGATGAGAGATGACTAAAAACAGGAGAAAATTTTTTGGCT 261
            ||  |||||||||||||  |   ||||||| ||||| |  |  |||| |||  |||||| |
Db 3062  TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAAACTTTGGGT 3118

Qy  262  GAGGAAGCAAAAACTTTGGCTGCGAGCCTTGCTTGAGAAGAGAGATAATAACTAAATTA 321
            |||||||||||||||  |  |||  |  |||||| |||||  | || ||  ||   ||| 
Db 3119  GAGGAAGCAAAAACTCTCAAGTGAAGTTGTGCTTGAAGAAGAGATTATGATGGAGTAATAA 3178

Qy  322  GAGATTAGAAGAAAATTAAGGAT---------GCAGTATCACACATAATAAAAGTTTCTA 370
            || || ||   ||||  ||||||          || |||||||||| ||||||||| | |
Db 3179  TTAAGTAGAGGTTAAATAAAGGATTTGAGTGTCAAAAAAACAAAAATTAATAAAAGTGTTGC 3238

Qy  371  CCTTTCTTAAAAGTGTAGC---TAATGTTGTGTTTTTAATTGGCTTTTTAGTAGCCTTTTTAT 427
            |||| |||||| ||| |||     | ||| || | | ||| | |   | ||||| |  ||| |
Db 3239  CTTTTTCTTATTAGGGTAGCTTGTGATGTTGTTGTTAGTATTGCCTATAGTAGTAGCCATTTGA 3298

Qy  428  TACACTTTAAATAAGTTGTGGCGACTTCAAT 456
            | ||| |||||||||||||  |||  | || 
Db 3299  CACATTAAATAAGTTGTTGTGACACATCATT 3327
```

```
Db   61  AGAAAAGCTTGTATCAGTGAGAAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGG 120
Qy  196  TGCCTTATGTTACTTAAGCCATGT 216
Db  121  TGCCTATGTTACTTAAGCCATGT 141
```

RESULT 8
US-10-072-809A-11
; Sequence 11, Application US20030217382A1
; Publication No. US20030217382A1
; GENERAL INFORMATION:
;   APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;   TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u...
;   TITLE OF INVENTION: therefor
;   FILE REFERENCE: 18-01
;   CURRENT APPLICATION NUMBER: US/10/072,809A
;   CURRENT FILING DATE: 2002-09-12
;   PRIOR APPLICATION NUMBER: USSN 60/267,271
;   PRIOR FILING DATE: 2001-02-08
;   NUMBER OF SEQ ID NOS: 61
;   SOFTWARE: Patentin version 3.0
;
;   SEQ ID NO 11
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Nicotiana alata
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(99)
US-10-072-809A-11

Query Match   18.3%; Score 99; DB 13; Length 99;
Best Local Similarity   100.0%; Pred. No. 3.5e-15;
Matches 99; Conservative   0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  217  GTGTTTGATGAGAAGATGACTAAAAACAGGAGCAAGAGCTGAAAATTTTGGGCTGAGAAGACAAAAACT 315
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GTGTTTGATGAGAAGATGACTAAAAACAGGAGCAAGAGCTGAAAATTTTGGGCTGAGAAGACAAAAACT 60
```

RESULT 9
US-10-072-809A-9
; Sequence 9, Application US20030217382A1
; Publication No. US20030217382A1
; GENERAL INFORMATION:
;   APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;   TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u...
;   TITLE OF INVENTION: therefor
;   FILE REFERENCE: 18-01
;   CURRENT APPLICATION NUMBER: US/10/072,809A
;   CURRENT FILING DATE: 2002-09-12
;   PRIOR APPLICATION NUMBER: USSN 60/267,271
;   PRIOR FILING DATE: 2001-02-08
;   NUMBER OF SEQ ID NOS: 61
;   SOFTWARE: Patentin version 3.0
;
;   SEQ ID NO 9
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Nicotiana alata
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(75)
US-10-072-809A-9

Query Match   13.9%; Score 75; DB 13; Length 75;
Best Local Similarity   100.0%; Pred. No. 3.2e-09;
Matches 75; Conservative   0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1  ATGGCTGCTCCTCGTTGTGCTTCATTGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTGCC 60
```

RESULT 10
US-10-312-841-1
Sequence 1, Application US/10312841
Publication No. US20030186277A1
GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match   8.9%;   Score 48;   DB 13;   Length 3673778;
Best Local Similarity 50.9%;   Pred. No. 4;
Matches 114;   Conservative 0;   Mismatches 110;   Indels 0;   Gaps 0;

Qy   1 ATGGCTGCTCCTCCTTGTGCTTCAGGCAGTGCATTGCCTATCTCTGGCAAGGATGGCTCCTTCGTTGTTCGCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   61 TATGAGGTGCAAGCT 75
           |||||||||||||||
Db   61 TATGAGGTGCAAGCT 75

Qy   299 AAGAGATAAATGGATAACTAATTAGAGATTAGAGAAATTAAGGATGCAGTATCACACATA 358
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66836 AAAGTATTAAAGAAAAATAAT:ATTATGAAAGAGAGTATTTGATTTATAAAATTTGAAAATA 66895

Qy   359 ATAAAGTTTCTACCTTTCCTTAAAAGTGTAGCTAATTGTTTAATTGGCTTTGGTTAGTA 418
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   66896 ATAATTTTTAATTTTGTTTTAATATGTA:GTAAAATTTTTTTGATATTTTTTTTTTTAATG 66955

Qy   419 GCCTTTTATTACACTTTAAATAAAGTGTGGCCACTCAATCCTTTGTGTGCAAATCTTGCACTAA 478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66956 GTGTAGTT:AATTTTTTTTTGTGAGTTCGGTTTGGGATTTAAATGATGTATTTTTTGATATG 67015

Qy   479 GTTTATTTGTGTACTTTTAAATGAAAATGAACCTTCATTAGTTCTTT 522
           |||||||||||||||||||||||||||||||||||||||||
Db   67016 GTTTTATTTTGTGTGTTTTATTTAAATTATTATTTGAATTGTAATT 67059

RESULT 11
US-10-178-213-82
Sequence 82, Application US/10178213
Publication No. US20030041348A1
GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 82
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Beta vulgaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)...(267)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (121)...(261)
US-10-178-213-82

Query Match   8.5%;   Score 46.2;   DB 15;   Length 516;
Best Local Similarity 57.1%;   Pred. No. 0.14;
Matches   84;   Conservative 0;   Mismatches 63;   Indels 0;   Gaps 0;

Qy   71 AAGCTAGAGAATGCAAAACAGAAAAGCAACACACATTTCCTGGAATATGCATTACCAAACCAC 130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116 AAGCAAGGACATGTATGACTCCAAGTCACCAGTTCAGGGGGAAATATGCGTTAGTAGTAGAA 175

Qy   131 CATGCAGAAAAGCTTGTATCAGTGAGAAAATTTACTGATGTCATTGTAGCAAAAACCTCA 190
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 ATTGTGAATCTGCTTGCCACACTGACAGAGATTTCCTGGAGGAACGGTGTCAAGGCCTTTGTA 235

Qy   191 GAAGGTGCCTATGTACTAAGGCCATGTG 217
           |||||||||||||||||||||||||||
Db   236 GAAGAATGCAGTGCACTAAGGCCTTGCG 262

RESULT 12
US-10-178-213-85
Sequence 85, Application US/10178213
Publication No. US20030041348A1
GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 85
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Beta vulgaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)...(267)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (121)...(261)
US-10-178-213-85

Query Match   8.2%;   Score 44.6;   DB 15;   Length 508;
Best Local Similarity 56.5%;   Pred. No. 0.35;
Matches   83;   Conservative 0;   Mismatches 64;   Indels 0;   Gaps 0;

Qy   71 AAGCTAGAGAATGCAAAAACAGAAAAGCAACAACACATTCCTGAATATGCATTACCAAACCAC 130
           |||||||||||||||||||||||||||||||||||||||||

RESULT 13
US-10-178-213-292
Sequence 292, Application US/10178213
Publication No. US20030041348A1
GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 292
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Beta vulgaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (62)...(301)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (152)...(292)
US-10-178-213-292

Query Match   8.2%; Score 44.6; DB 15; Length 578;
Best Local Similarity   56.5%; Pred. No. 0.37;
Matches   83; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy   71 AAGCTAGAGAATGCAAAACAGAAAGCAACACATTCCTGGAATATGCATTACCAAACCAC 130
          ||| |||||| |||||| || || | || ||||| | || | || ||| | |
Db   147 AAGCAGAACTTGTGGAACTCCAAGCCAGAGGTTCAGCGGGATTTATGTTGTTAGAAGAGAA 206

Qy   131 CATCGAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTGAGCAAAATCCTCA 190
          | ||| ||| ||| | | || ||| |||| || ||| || || | |
Db   207 ATTGTGAATCAGTTTGCAACAGTGAAGGAATTTCCTGATGGAAGTTGTCAAGGCGCACGCA 266

Qy   191 GAAGGTGCCTATGTACTAAGCCATGTG 217
          ||||| | ||| || ||||||||||
Db   267 GAAGANGTATTTGCAACAGGCCTTGCG 293

RESULT 14
US-10-311-455-1685
Sequence 1685, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ...

; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-07-02
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1685
; LENGTH: 13326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1685

Query Match   8.2%; Score 44.6; DB 13; Length 13326;
Best Local Similarity   47.3%; Pred. No. 1.8;
Matches   134; Conservative   0; Mismatches 149; Indels   0; Gaps   0;

Qy   254 TTTTCGGCTGAGGAGGCACAAAAACTTGGCGTGCCAGCTTTGCTTTGGAAGAGAAGGAATAAGGATA 313
          ||| ||||| || || ||||| | |||||| ||| || |||| | ||||| || || ||||
Db 10474 TTATATTTTAAGGATAAATAAATTTGTTGTTGTTTAAGTGTTGTTATTGGTGGTTTGTT 10533

Qy   314 ACTAATTAGGATTAGGAGAGAAAATTAAGGANGCAGTATCACACATAATAAAGTTCTACCT 373
          || ||||||| |||| || ||| |||||| || ||| |||| |||| ||||| |
Db 10534 ACGGTAGTAATAGAAAAATTAATTTAAGAAGTAAATGAGATAATTTATATGAGTGAATAGGT 10593

Qy   374 TTCTTAAAAGTGTAGCTAATGTTGTGTTTTAATTGGCTTTTAGTTAGCCTTTTATTACACT 433
          ||| | ||| | || || | ||| | | | | || || |||| || |||| || |
Db 10594 ATGTAGTAAGTAATAAGTAGTATTAGGTTTTGAATTGTTGTTTGTGTATTTTATAATGAGA 10653

Qy   434 TTAAATAAGTGTGGCGCACTTCAAATCCTTTGTGCAATCTTGCACTAAGTTTATTTGTGTACT 493
          |||||| |||| || |||| || | ||||| || ||||| ||||| | |||||||
Db 10654 TTGTTTGGTTTTAGTGATGGAGAGTATGAGGAAGTGGGTTAAAGTTTTTTTTTTAGGGTAT 10713

Qy   494 TTTAATGAAAATGACCTTCTATGGCTCTTTGGTTAAAAAAAAAAA 536
          || ||| ||||| | | | | ||||| || |||||| ||||
Db 10714 TTTTTATTTTATTATTTTGATAAAGGTTATTAATTAATTTGTAAAAA 10756

RESULT 15
US-09-294-093B-4663
Sequence 4663, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4663
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 70035475SH1
; NAME/KEY: unsure
; LOCATION: 14, 100
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4663

Query Match   8.0%; Score 43.4; DB 9; Length 283;

Best Local Similarity   54.8%; Pred. No. 0.52; Indels   0; Gaps   0;
Matches 86; Conservative   0; Mismatches   71;

```
Qy    59 CCTATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAGCAACACATTTCCTGGAATATGCA 118
         || ||    || |  | | ||  |  || ||||||||   | | ||  ||||||| ||||
Db   111 CCGATATGGCACAGGCAAGGGAATGCCGAGAAGTACACAGTGAGCGATTTGTTGGGGCATGCA 170

Qy   119 TTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTA 178
         |   ||  |||     ||     |  |  | |   || ||  |   |  |  || | |
Db   171 TGATCGCAGACAACTGCGCCCAATGTGTTGCCCGCGGGTGAGGGCTTCTTGGGCCGGCAGTGCA 230

Qy   179 GCAAAATCCTCAGAAGGTGCCTTATGTACTAAGCCATG 215
         || |   |    |  | || | || ||| |  ||||||
Db   231 GCACCTTCCGCCCGCCCTGCATCTGCACTAGGCAGTG 267
```

Search completed: January 28, 2004, 11:23:53
Job time : 1177 secs

OM nucleic - nucleic search, using sw model

Copyright (c) GenCore version 5.1.6
1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 08:44:46 ; Search time 56 Seconds
(without alignments)
4264.076 Million cell updates/sec

Title: US-10-072-809B-17
Perfect score: 541
Sequence: 1 atggctcgctccttgtgctt..........tggttaaaaaaaaaaaaaa 541

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	189.8	35.1	564	3	US-08-984-320-1	Sequence 1, Appli
2	189.8	35.1	564	1	US-08-487-087A-1	Sequence 1, Appli
3	189.8	35.1	564	1	US-08-397-653B-1	Sequence 1, Appli
4	189.8	35.1	564	6	US-517-5095-1	Patent No.5175095
5	147.4	27.2	3528	3	US-08-984-320-2	Sequence 2, Appli
6	147.4	27.2	3528	1	US-08-487-087A-2	Sequence 2, Appli
7	147.4	27.2	4383	4	US-08-397-653B-2	Sequence 2, Appli
8	147.4	27.2	4383	6	US-5175095-4	Patent No.5175095
9	147.4	27.2	4383	5	US-5173307-1	Patent No.5177307
10	95.6	17.7	506	4	US-09-442-631-1	Sequence 1, Appl
11	45.4	8.4	7218	1	US-08-232-463-14	Sequence 14, Appl
12	42	7.8	1692	1	US-09-601-198-63	Sequence 63, Appl
13	38	7.0	53332	1	US-09-801-861-3	Sequence 3, Appl
14	37.2	6.9	1738	2	US-08-472-028A-3	Sequence 3, Appl
15	37.2	6.9	1738	3	US-08-808-323-3	Sequence 3, Appl
16	37.2	6.9	1738	3	US-08-808-931-3	Sequence 3, Appl
17	37.2	6.9	1738	3	US-09-050-603A-3	Sequence 3, Appl
18	37.2	6.9	1738	3	US-09-102-420B-3	Sequence 3, Appl
19	37.2	6.9	1738	3	US-09-071-296-3	Sequence 3, Appl
20	37.2	6.9	1738	3	US-09-196-094-3	Sequence 3, Appl
21	37.2	6.9	1738	3	US-09-015-683-3	Sequence 3, Appl
22	37.2	6.9	1738	4	US-09-191-998-3	Sequence 3, Appl
23	37.2	6.9	1738	4	US-08-497-698-3	Sequence 3, Appl
24	37.2	6.9	19124	4	US-08-487-826B-13	Sequence 13, Appl
25	36.8	6.8	225	4	US-08-442-631-3	Sequence 3, Appl
26	36.8	6.8	10607	1	US-08-078-090-3	Sequence 3, Appl
27	36.2	6.7	3000	2	US-08-680-395-1	Sequence 1, Appl
c 28	36	6.7	1939	1	US-07-715-751B-2	Sequence 2, Appli
c 29	35.4	6.5	19124	2	US-08-487-826B-13	Sequence 13, Appl
c 30	35	6.5	6701	4	US-09-611-559A-1	Sequence 1, Appli
c 31	35	6.5	580073	4	US-08-545-528D-1	Sequence 1, Appli
c 32	35	6.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c 33	35	6.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c 34	34.8	6.4	2943	4	US-09-503-922-2	Sequence 2, Appli
c 35	34.6	6.4	2028	2	US-08-933-750C-75	Sequence 75, Appl
c 36	34.6	6.4	2028	2	US-09-234-613-75	Sequence 75, Appl
c 37	34.6	6.4	2028	1	US-08-184-252A-1	Sequence 1, Appli
c 38	34.6	6.4	3000	5	PCT-US95-00601-1	Sequence 1, Appli
c 39	34.6	6.4	162450	4	US-09-345-882-1	Sequence 1, Appli
c 40	34.4	6.4	1786	1	US-07-920-430-19	Sequence 19, Appli
c 41	34.4	6.4	1786	1	US-08-066-299-9	Sequence 9, Appli
c 42	34.4	6.4	1786	1	US-08-265-047-1	Sequence 1, Appli
c 43	34.4	6.4	1786	1	US-08-251-464-19	Sequence 19, Appli
c 44	34.4	6.4	1786	5	US-08-926-522-20	Sequence 20, Appl
c 45	34.4	6.4	1786	5	PCT-US92-03364-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-08-984-320-1
Sequence 1, Application US/08984320
Patent No. 6222097
GENERAL INFORMATION:
APPLICANT: McBride, Kevin E.
APPLICANT: Stalker, David M.
TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,320
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/487,087
FILING DATE: 07-JUN-95
APPLICATION NUMBER: USSN 07/998,158
FILING DATE: 29-DEC-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/554,195
FILING DATE: 17-JUL-90
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/382,518
FILING DATE: 19-JUL-89
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
REFERENCE/DOCKET NUMBER: CGNE 91-1
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

```
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
US-08-984-320-1

  Query Match   35.1%; Score 189.8; DB 3; Length 564;
  Best Local Similarity   68.9%; Pred. No. 7.9e-42;
  Matches 326; Conservative   0; Mismatches 127; Indels   20; Gaps   4;

Qy   1 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC 60
              |||||||| ||| | | ||||||||||||| ||||||||| ||||||||||||| ||
Db   40 ATGGCTCGTTCCATTTTCTTCATGGCATTTTTGGTCTTGGCAATGATGCTCTTTGTTACC 99

Qy   61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTTCCTGGAATATGC 117
              |||||||| ||||| || |||||| || |||| | || || || ||| ||||
Db   100 TATGAGGTAGAAGCTCAGCAAATTTGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGT 159

Qy   118 ATTACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
               ||| | || |||| |||||| |||||||| ||||||||||||| ||| ||||||
Db   160 TTTATGGACTCATCATGTAGAAAATATTGTATCAAAGAGAAATTTACTGGTGGACATTGT 219

Qy   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGACT 237
              |||||| ||| || | ||| ||||| |||||||||||||| || ||| || || |
Db   220 AGCAAACTCCAAAGGAAGTGTCTATGCACTAAGCCATGTGTATT---TGACAAAATCTCA 276

Qy   238 AAAACAGGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAA 297
              | || | || ||||| |||||||||||||||| | || || | ||||||||
Db   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336

Qy   298 GAAGAGATAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGAT-----------GCA 346
              |||||||| ||| | |||| | | | || ||||||| |
Db   337 GAAGAGATTATGATGGAGTAATAATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 396

Qy   347 GTATCACACATAATAAAGTTTCTACCTTTCTTAAAAGTGTAGC---TAATGTTGTGTTTT 403
                | || | ||||||||| | | ||||||| || ||||| | ||||||||||
Db   397 AAAACAAAATTAATAAAGTGTTGCCTTTTCTTATTAGGGTAGCTTGTGATGTTGTGTTAG 456

Qy   404 AATTGGCTTTTAGTAGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456
               |||||| | |||||||| ||| ||| | | | |||| ||| ||| |
Db   457 TATTGGCCTATAGTAGCCATTTGACACATTAAATAAGTTTGTGACACATCATT 509

RESULT 2
US-08-487-087A-1
; Sequence 1, Application US/08487087A
; Patent No. 6268546
; GENERAL INFORMATION:
;   APPLICANT: McBride, Kevin E.
;   APPLICANT: Stalker, David M.
;   TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calgene, Inc.
;   STREET: 1920 Fifth Street
;   CITY: Davis
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 95616
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
;   COMPUTER: Apple Macintosh
;   OPERATING SYSTEM: Macintosh 7.1
;   SOFTWARE: Microsoft Word 5.1 (a)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/487,087A
;   FILING DATE: 07-JUN-95
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 07/998,158
;   FILING DATE: 29-DEC-92
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 07/554,195
;   FILING DATE: 17-JUL-90
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 07/382,518
;   FILING DATE: 19-JUL-89
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Donna E. Scherer
;   REGISTRATION NUMBER: 34,719
;   NAME: Carl J. Schwedler
;   REGISTRATION NUMBER: 36,924
;   REFERENCE/DOCKET NUMBER: CGNE 91-1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (916) 753-6313
;   TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 564 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
US-08-487-087A-1

  Query Match   35.1%; Score 189.8; DB 3; Length 564;
  Best Local Similarity   68.9%; Pred. No. 7.9e-42;
  Matches 326; Conservative   0; Mismatches 127; Indels   20; Gaps   4;

Qy   1 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC 60
              |||||||| ||| | | ||||||||||||| ||||||||| ||||||||||||| ||
Db   40 ATGGCTCGTTCCATTTTCTTCATGGCATTTTTGGTCTTGGCAATGATGCTCTTTGTTACC 99

Qy   61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTTCCTGGAATATGC 117
              |||||||| ||||| || |||||| || |||| | || || || ||| ||||
Db   100 TATGAGGTAGAAGCTCAGCAAATTTGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGT 159

Qy   118 ATTACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
               ||| | || |||| |||||| |||||||| ||||||||||||| ||| ||||||
Db   160 TTTATGGACTCATCATGTAGAAAATATTGTATCAAAGAGAAATTTACTGGTGGACATTGT 219

Qy   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGACT 237
              |||||| ||| || | ||| ||||| |||||||||||||| || ||| || || |
Db   220 AGCAAACTCCAAAGGAAGTGTCTATGCACTAAGCCATGTGTATT---TGACAAAATCTCA 276

Qy   238 AAAACAGGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAA 297
              | || | || ||||| |||||||||||||||| | || || | ||||||||
Db   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336

Qy   298 GAAGAGATAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGAT-----------GCA 346
              |||||||| ||| | |||| | | | || ||||||| |
Db   337 GAAGAGATTATGATGGAGTAATAATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 396

Qy   347 GTATCACACATAATAAAGTTTCTACCTTTCTTAAAAGTGTAGC---TAATGTTGTGTTTT 403
                | || | ||||||||| | | ||||||| || ||||| | ||||||||||
Db   397 AAAACAAAATTAATAAAGTGTTGCCTTTTCTTATTAGGGTAGCTTGTGATGTTGTGTTAG 456

Qy   404 AATTGGCTTTTAGTAGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456
               |||||| | |||||||| ||| ||| | | | |||| ||| ||| |
Db   457 TATTGGCCTATAGTAGCCATTTGACACATTAAATAAGTTTGTGACACATCATT 509

RESULT 3
US-08-397-653B-1
; Sequence 1, Application US/08397653B
; Patent No. 6329570
; GENERAL INFORMATION:
;   APPLICANT: Martineau, Belinda
;   TITLE OF INVENTION: COTTON MODIFICATION USING
;   TITLE OF INVENTION: OVARY-TISSUE TRANSCRIPTIONAL
;   TITLE OF INVENTION: FACTORS
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calgene, Inc.
;   STREET: 1920 Fifth Street
;   CITY: Davis
;   STATE: CA
;   COUNTRY: USA
```

;     ZIP: 95616
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
;       COMPUTER: Apple Macintosh
;       OPERATING SYSTEM: Macintosh 7.0
;       SOFTWARE: Microsoft Word 5.1a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/397,653B
;       FILING DATE: 28-FEB-1995
;       CLASSIFICATION: 800
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Elizabeth Lassen
;         REGISTRATION NUMBER: 31,845
;       NAME: Donna E. Scherer
;         REGISTRATION NUMBER: 34,719
;       NAME: Carl J. Schwedler
;         REGISTRATION NUMBER: 36,924
;       REFERENCE/DOCKET NUMBER: CGNE 112
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (916) 753-6313
;       TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 564 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
; US-08-397-653B-1

Query Match    35.1%;  Score 189.8;  DB 4;  Length 564;
Best Local Similarity  68.9%;  Pred. No. 7.9e-42;
Matches 326;  Conservative  0;  Mismatches 127;  Indels  20;  Gaps  4;

Qy      1 ATGGCTCGCTCCTTGTGCTTCATGGCCATTTGCTACTTGCCAAGGAATGCTCTTTCTTGCC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 ATGGCTCGTTCCATTTTCTTCATCGGCATTTTTGGTCTTTGGCAAGTGCTCTTTGTTTGCC 99

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAAGCAACACATTTCCTGGAATATGC 117
            |||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db    100 TATGAGGTAGAAGCTCAGCAAAATTTGCAAAGCACCAAAACTTTCCCAGGATTATGT 159

Qy    118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 TTTATGGACTCATCATGTAGAAAATAATTGTAATCAAAGAGAAATTTACTGTGGACATTGT 219

Qy    178 AGCAAAAATCCTCAGAAGGTGCCTAGTGACTAAGCCATGGTGTTTGATGAGAAGATGACT 237
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 AGCAAACTCCAAAGGAAGTGTCTAATGCACTAAGCCATGTGTTATT---TGACAAAAAATCTCA 276

Qy    238 AAAACAGAGACGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCCAGCTTTGCTTGAA 297
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277 AGTGAAGTTAAAGCAACTTTGGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGGTGCTTGAA 336

Qy    298 GAAGAGATAAATGGATAACTAATTAATTAGAAGATTAGAAGAAAAATTAAGGAT-------GCA 346
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337 GAAGAGATTATGATGGAGTAAATAATAACGGATTTTGAGGTGTCAAAA 396

Qy    347 GTATCACACATAATAAAGTTCTACCTTTCTTACCTTTCTTAAAAGTGTAGC- 403
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    397 AAAACAAAATTAATAAAGTGTTGGCCTTTCTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456

Qy    404 AATTGGCTTTTAGTAGCGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 TATTGGGCCTATAGTAGCAATTGACACAATAAATAAGTTTGTGACACATCATT 509
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/07/554,195
;       FILING DATE: 17-JUL-1990
; SEQ ID NO:1:
;     LENGTH: 564
; 5175095-1

Query Match    35.1%;  Score 189.8;  DB 6;  Length 564;
Best Local Similarity  68.9%;  Pred. No. 7.9e-42;
Matches 326;  Conservative  0;  Mismatches 127;  Indels  20;  Gaps  4;

Qy      1 ATGGCTCGCTCCTTGTGCTTCATGGCCATTTGCTACTTGCCAAGGAATGCTCTTTCTTGCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 ATGGCTCGTTCCATTTTCTTCATCGGCATTTTTGGTCTTTGGCAAGTGCTCTTTGTTTGCC 99

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAAGCAACACATTTCCTGGAATATGC 117
            ||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db    100 TATGAGGTAGAAGCTCAGCAAAATTTGCAAAGCACCAAAACTTTCCCAGGATTATGT 159

Qy    118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 TTTATGGACTCATCATGTAGAAAATAATTGTAATCAAAGAGAAATTTACTGTGGACATTGT 219

Qy    178 AGCAAAAATCCTCAGAAGGTGCCTAGTGACTAAGCCATGGTGTTTGATGAGAAGATGACT 237
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 AGCAAACTCCAAAGGAAGTGTCTAATGCACTAAGCCATGTGTTATT---TGACAAAAAATCTCA 276

Qy    238 AAAACAGAGACGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCCAGCTTTGCTTGAA 297
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    277 AGTGAAGTTAAAGCAACTTTGGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGGTGCTTGAA 336

Qy    298 GAAGAGATAAATGGATAACTAATTAATTAGAAGATTAGAAGAAAAATTAAGGAT-------GCA 346
            |||||||||||||||||||||||||||||||||||||||||||||
Db    337 GAAGAGATTATGATGGAGTAAATAATAACGGATTTTGAGGTGTCAAAA 396

Qy    347 GTATCACACATAATAAAGTTCTACCTTTCTTACCTTTCTTAAAAGTGTAGC- 403
            ||||||||||||||||||||||||||||||||||||||||||||
Db    397 AAAACAAAATTAATAAAGTGTTGGCCTTTCTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456

Qy    404 AATTGGCTTTTAGTAGCGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    457 TATTGGGCCTATAGTAGCAATTGACACAATAAATAAGTTTGTGACACATCATT 509
```

RESULT 4
;Patent No.5175095
;   APPLICANT: Martineau, Belinda M.;Houck, Catherine M.
;   TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS
;   NUMBER OF SEQUENCES: 9

RESULT 5
US-08-984-320-2
; Sequence 2, Application US/08984320
; Patent No. 6222097
; GENERAL INFORMATION:
;   APPLICANT: McBride, Kevin E.
;   APPLICANT: Stalker, David M.
;   TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
;   NUMBER OF SEQUENCES: 6
;   SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/984,320
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calgene, Inc.
;   STREET: 1920 Fifth Street
;   CITY: Davis
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 95616
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
;   COMPUTER: Apple Macintosh
;   OPERATING SYSTEM: Macintosh 7.1
;   SOFTWARE: Microsoft Word 5.1 (a)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/984,320
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/487,087
;   FILING DATE: 07-JUN-95
;   APPLICATION NUMBER: USSN 07/998,158

RESULT 6
US-08-487-087A-2
; Sequence 2, Application US/08487087A

GENERAL INFORMATION:
; APPLICANT: McBride, Kevin E.
; APPLICANT: Stalker, David M.
; TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA

FILING DATE: 29-DEC-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/554,195
FILING DATE: 17-JUL-90
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/382,518
FILING DATE: 19-JUL-89
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 91-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-984-320-2

Query Match　　27.2%;　Score 147.4;　DB 3;　Length 3528;
Best Local Similarity　67.1%;　Pred. No. 3.1e-30;
Matches 261;　Conservative　0;　Mismatches 111;　Indels　17;　Gaps　3;

```
Qy    82 TGCAAAACAGAAAGCAACACAACATTTCCTGGAATATGCAATTACCAAACCACCATGCAGAAAA 141
          ||||||      |||||||||||| |||||||||||||||||| ||| || ||||||| |||||
Db  2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTTATTAGGACTCATCATGTAGAGAAA 3001

Qy   142 GCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
          |||||||    ||||||||||||||||||||| ||||||||||    |||||||||| ||
Db  3002 TATTGTATCAAAGAGAGAAATTTACTGGTGGACATTGTAGCAAAACTCCAAAGGAAGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTTGTTTGATGAGAAGATGACTAAAAACAGGAGGCTGAAATTTTGGCT 261
          |||||||||||||||| |||| |||||||||  |||| |    ||||||||||||||||||||
Db  3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAAACTTTGGGCTGCAGCTTTGCTTGAAGAAGAGATAATGGATAATAACTAATTA 321
          ||||||||||||| ||||  |||||||||||||||||||||| |||||||||||||  |||||||
Db  3119 GAGGAAGCAAAAAACTCTAAGTGAAGTTGTGCTTGAAGAAGAGATTATGAGGAGTAATAA 3178

Qy   322 GAGATTAGAAGAAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTGTTCTA 370
          |||||||||||||||||||||||          ||||||||| | | |||   |||| |||
Db  3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAAAAAACAAAAATTAATAAAAGTGTTGC 3238

Qy   202 TGTACTAAGCCATGTGTTGTTTGATGAGAAGATGACTAAAAACAGGAGGCTGAAATTTTGGCT 261
          ...
Db  3062 ...

Qy   371 CCTTTCTTAAAAGTGTAGC---TAATGTTGTGTTTTAATTGGCTTTTTTAGTAGCCTTTTTAT 427
          |||||  ||||||||||||   ||||| |||| ||||||| ||||||||||||||||||   ||
Db  3239 CTTTTCTTTATTAGGGTAGCCTTGTGATGTTGTTAGTATTAGGCCTATAGTAGCCCATTTGA 3298

Qy   428 TACACTTTAAATAAGTTGTGGGCACTTCAAT 456
          |||||||||||||| ||||||||||||||||
Db  3299 CACATTAAATAAGTTTGTGACACATCATT 3327
```

RESULT 7
US-08-397-653B-2
; Sequence 2, Application US/08397653B

ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,087A
FILING DATE: 07-JUN-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/998,158
FILING DATE: 29-DEC-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/554,195
FILING DATE: 17-JUL-90
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/382,518
FILING DATE: 19-JUL-89
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 91-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-487-087A-2

Query Match　　27.2%;　Score 147.4;　DB 3;　Length 3528;
Best Local Similarity　67.1%;　Pred. No. 3.1e-30;
Matches 261;　Conservative　0;　Mismatches 111;　Indels　17;　Gaps　3;

```
Qy    82 TGCAAAACAGAAAGCAACACAACATTTCCTGGAATATGCAATTACCAAACCACCATGCAGAAAA 141
          ||||||      |||||||||||| |||||||||||||||||| ||| || ||||||| |||||
Db  2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTTATTAGGACTCATCATGTAGAGAAA 3001

Qy   142 GCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
          |||||||    ||||||||||||||||||||| ||||||||||    |||||||||| ||
Db  3002 TATTGTATCAAAGAGAGAAATTTACTGGTGGACATTGTAGCAAAACTCCAAAGGAAGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTTGTTTGATGAGAAGATGACTAAAAACAGGAGGCTGAAATTTTGGCT 261
          |||||||||||||||| |||| |||||||||  |||| |    ||||||||||||||||||||
Db  3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAAACTTTGGGCTGCAGCTTTGCTTGAAGAAGAGATAATGGATAATAACTAATTA 321
          ||||||||||||| ||||  |||||||||||||||||||||| |||||||||||||  |||||||
Db  3119 GAGGAAGCAAAAAACTCTAAGTGAAGTTGTGCTTGAAGAAGAGATTATGAGGAGTAATAA 3178

Qy   322 GAGATTAGAAGAAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTGTTCTA 370
          |||||||||||||||||||||||          ||||||||| | | |||   |||| |||
Db  3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAAAAAACAAAAATTAATAAAAGTGTTGC 3238

Qy   371 CCTTTCTTAAAAGTGTAGC---TAATGTTGTGTTTTAATTGGCTTTTTTAGTAGCCTTTTTAT 427
          |||||  ||||||||||||   ||||| |||| ||||||| ||||||||||||||||||   ||
Db  3239 CTTTTCTTTATTAGGGTAGCCTTGTGATGTTGTTAGTATTAGGCCTATAGTAGCCCATTTGA 3298

Qy   428 TACACTTTAAATAAGTTGTGGGCACTTCAAT 456
          |||||||||||||| ||||||||||||||||
Db  3299 CACATTAAATAAGTTTGTGACACATCATT 3327
```

Patent No. 6329570
GENERAL INFORMATION:
 APPLICANT: Martineau, Belinda
 TITLE OF INVENTION: COTTON MODIFICATION USING
 TITLE OF INVENTION: OVARY-TISSUE TRANSCRIPTIONAL
 TITLE OF INVENTION: FACTORS
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: Calgene, Inc.
  STREET: 1920 Fifth Street
  CITY: Davis
  STATE: CA
  COUNTRY: USA
  ZIP: 95616
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
  COMPUTER: Apple Macintosh
  OPERATING SYSTEM: Macintosh 7.0
  SOFTWARE: Microsoft Word 5.1a
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/397,653B
  FILING DATE: 28-FEB-1995
  CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
  NAME: Elizabeth Lassen
  REGISTRATION NUMBER: 31,845
  NAME: Donna E. Scherer
  REGISTRATION NUMBER: 34,719
  NAME: Carl J. Schwedler
  REGISTRATION NUMBER: 36,924
  REFERENCE/DOCKET NUMBER: CGNE 112
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (916) 753-6313
  TELEFAX: (916) 753-1510
 INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
   LENGTH: 4383 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
   MOLECULE TYPE: cDNA to mRNA
US-08-397-653B-2

Query Match 27.2%; Score 147.4; DB 4; Length 4383;
Best Local Similarity 67.1%; Pred. No. 3.3e-30;
Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

```
Qy    82 TGCAAAAACAGAAAGCAACACACATTTCCTGGAATATGCATTACCAAAACCACCATGCAGAAAA    141
         ||||| ||  |||||| ||||  |||||||||||||| |||||||||||||| |||| |||||||
Db  2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTATGGACTCATCATGTAGAAAA        3001

Qy   142 GCTTGTATCACGTGAGAGAAATTTACTCATGGTGTCATTGTGAGCAAAAATCCTCAGAAGGTGCCTA   201
         |||||| | |  |||||||||||||||| |||||||||| |||||||||||||||||| |||| |||||
Db  3002 TATTGTATCAAAGAGAGAAATTTACTGGTGGACACATTGTAGCAAACTCCAAAGGAAGTGTCTA     3061

Qy   202 TGTACTAAGCCATGTGTGTGTTTGATGAAGAAGATGACTAAAACAGGAGCTGAAATTTTGGCT     261
         |||||| ||||| |||||||  |||| ||||||| || |||||||||  ||||||||||| |||||
Db  3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT        3118

Qy   262 GAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAAGAAGGATAATGGATAACTAATTA    321
         ||||||||||||| |||||||||| |||| ||| |||  ||||||| |||||||| ||  ||||
Db  3119 GAGGAAGCAAAAACTCAAGTGAAGTTGTGCTTGAAGAAGAGGATTATGGATGGTAATAA       3178

Qy   322 GAGATTAGAAGAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTTTCTA    370
         |||||||| ||||| |||||||          |||||||||||| |||||| ||||||||
Db  3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAAAACAAAAATTAATAAAAGTGTTGC    3238

Qy   371 CCTTTCTTTAAAACTGTAGC----------TAATGTTGTTGTTTTTAATTAGTAGCCCTTTTAT    427
         ||||||||||||| |||||||          |||||||||| ||||||||||||||||||| ||
Db  3239 CTTTTCTTATTAGGGTAGCTTGTGAGTGTTGTTAGTATTGGCCTATAGTAGTAGCCATTTGA    3298

Qy   428 TACACTTTAAATAAGTGTGGCACTCAAT    456
         ||||| ||| ||||| ||| ||| ||||
Db  3299 CACATTAAATAAGTTTGTGACACATCATT    3327
```

RESULT 8
5175095-4
;Patent No. 5175095
;APPLICANT: Martineau, Belinda M.;Houck, Catherine M.
;TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS
;NUMBER OF SEQUENCES: 9
;CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/554,195
; FILING DATE: 17-JUL-1990
;SEQ ID NO:4:
; LENGTH: 4383
5175095-4

Query Match 27.2%; Score 147.4; DB 6; Length 4383;
Best Local Similarity 67.1%; Pred. No. 3.3e-30;
Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

```
Qy    82 TGCAAAAACAGAAAGCAACACACATTTCCTGGAATATGCATTACCAAAACCACCATGCAGAAAA    141
         ||||| ||  |||||| ||||  |||||||||||||| |||||||||||||| |||| |||||||
Db  2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTATGGACTCATCATGTAGAAAA        3001

Qy   142 GCTTGTATCACGTGAGAGAAATTTACTCATGGTGTCATTGTGAGCAAAAATCCTCAGAAGGTGCCTA   201
         |||||| | |  |||||||||||||||| |||||||||| |||||||||||||||||| |||| |||||
Db  3002 TATTGTATCAAAGAGAGAAATTTACTGGTGGACACATTGTAGCAAACTCCAAAGGAAGTGTCTA     3061

Qy   202 TGTACTAAGCCATGTGTGTGTTTGATGAAGAAGATGACTAAAACAGGAGCTGAAATTTTGGCT     261
         |||||| ||||| |||||||  |||| ||||||| || |||||||||  ||||||||||| |||||
Db  3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT        3118

Qy   262 GAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAAGAAGGATAATGGATAACTAATTA    321
         ||||||||||||| |||||||||| |||| ||| |||  ||||||| |||||||| ||  ||||
Db  3119 GAGGAAGCAAAAACTCAAGTGAAGTTGTGCTTGAAGAAGAGGATTATGGATGGTAATAA       3178

Qy   322 GAGATTAGAAGAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTTTCTA    370
         |||||||| ||||| |||||||          |||||||||||| |||||| ||||||||
Db  3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAAAACAAAAATTAATAAAAGTGTTGC    3238

Qy   371 CCTTTCTTTAAAACTGTAGC----------TAATGTTGTTGTTTTTAATTAGTAGCCCTTTTAT    427
         ||||||||||||| |||||||          |||||||||| ||||||||||||||||||| ||
Db  3239 CTTTTCTTATTAGGGTAGCTTGTGAGTGTTGTTAGTATTGGCCTATAGTAGTAGCCATTTGA    3298

Qy   428 TACACTTTAAATAAGTGTGGCACTCAAT    456
         ||||| ||| ||||| ||| ||| ||||
Db  3299 CACATTAAATAAGTTTGTGACACATCATT    3327
```

RESULT 9
5177307-1
;Patent No. 5177307
;APPLICANT: HOUCK, CATHERINE M.;PEAR, JULIE R.;MARTINEAU,
;BELINDA M.;HIATT, WILLIAM
;TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;MODULATIONS OF ENDOGENOUS CYTOKININ LEVELS
;NUMBER OF SEQUENCES: 6
;CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/554,196
; FILING DATE: 17-JUL-1990
;PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 382,802
; FILING DATE: 19-JUL-1989
; APPLICATION NUMBER: 188,361
; FILING DATE: 29-APR-1988
; APPLICATION NUMBER: 168,190
; FILING DATE: 15-MAR-1988
; APPLICATION NUMBER: 54,369
; FILING DATE: 26-MAY-1987
;SEQ ID NO:1:
; LENGTH: 4383
5177307-1

```
Query Match         27.2%;  Score 147.4;  DB 6;  Length 4383;
Best Local Similarity   67.1%;  Pred. No. 3.3e-30;
Matches 261;  Conservative  0;  Mismatches 111;  Indels 17;  Gaps 3;

Qy   82 TGCAAAACAGAAAGCAACACATTTCCTGGAATATGCCATTACCAACCACCATGCGAGAAAA 141
        ||||||  |||  |||| ||| ||||||| ||| ||| || ||  |  ||| || ||||||
Db 2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTTTATGGACTCATCATGTAGAAAA 3001

Qy  142 GCTTGTATCAGTGAGAGAAATTTACTGATTGTTGTAGCAAAAATCCTCAGAAGGTGCCTA 201
        || ||||||||| |||||||||||||| |||| || ||||||||| ||   | |||| |||
Db 3002 TATTGTATCAAAGAGAAATTTACTGGTGGACATTGTTGAGCAAAAATCCAAAGAAGTGTCTA 3061

Qy  202 TGTACTTAAGCCCATTGTGTTGATGAGAGATGACTAAAACAGGAGGCTGAAAATTTTGGCT 261
        || |||||||| ||||| |||  |    | |  |||||| |||    ||||| |||||||| 
Db 3062 TGCACTTAAGCCATTGTTGTATT---TGACAAAAATCTCAAGCTGAAGTTAAAGCAACTTTGGGT 3118

Qy  262 GAGGAAGCAAAAACTTTGGCTGCACGTTTGCTTGATTGAAGAGATAAATGGATAACTAAATTA 321
        ||||||||||||||| | |||||| ||||||||||| | |||||||| ||| ||| |||| ||
Db 3119 GAGGAAGCAAAAACTCTAAGCGAAGTTGTGCTTGAAGAGATTATGATGGAGTAATAA 3178

Qy  322 GAGATTAGAAGAAAATTAAGGAT---------GCAGTATCACACATAATAAAGTTTCTA 370
        || |||| | ||||| |||||||         |||| | | |||| |||||||||| ||
Db 3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGGTCAAAAAAAACAAAATTAATAAAAGTGTTGGC 3238

Qy  371 CCTTTCTTTAAAAGTGTAGC----TAAAGTTGTGTTTTTTAATTGGCTTTTTTTAT 427
        |||||||| |||||||||||    || |||||||||| | | |||||| |||||||
Db 3239 CTTTTCTTATTAGGGTAGCTTTGTGATGTTGTGTTAGTTATTGGCCTATAGTAGGCCATTTGA 3298

Qy  428 TACACTTTAAATAAGTGTGGGCACTTCAAT 456
        ||||||| |||||||||| | ||| |||||
Db 3299 CACATTAAATAAGTTGTTGACACATCATT 3327
```

RESULT 10

US-09-442-631-1
; Sequence 1, Application US/09442631
; Patent No. 6300489
; GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-442-631-1

```
Query Match         17.7%;  Score 95.6;  DB 4;  Length 506;
Best Local Similarity   56.9%;  Pred. No. 1.2e-16;
Matches 312;  Conservative  0;  Mismatches 174;  Indels 62;  Gaps 5;

Qy    1 ATGGCTCGCTCCTTGTTGCTTCATGGCATTGGTATCTTGGCAAGGATGGCTCCTTCTTGTTGCC 60
        ||||||||| || | |||  ||||||||||  | ||| |||||| || | | ||||| ||||||
Db   10 ATGGCTCGTTCCATTTACTTTCATGGCATTCTTCATGCTTGTCTTGGCAAATGAAGACTCTTCTTGTTGGCT 69

Qy   61 TATGAGGTGCCAAGCTAGAGAATGCAAAAACAGAGAAAGCAACACATTTCCTGGAATATGCCATT 120
        |||| ||||| || | ||| | | |||||||||||||| |    | |  |   |  || | ||| 
Db   70 TATGGGGTGCAAGGCAAGGAAAAATTTGCTGTAAAACAGAGAAATTTGCTGTAAAGAGCTCACAAAACCTGTTAAAATGTTCT 129

Qy  121 ACCAAACCACCATGCGAGAAAAGCTTGTAT---CAGTGAGAGAAATTTACTGATTGCTCATTGT 177
        | | | | |||| | | ||||| | | |||   || ||| |||||| |  |  || | |  |||
Db  130 AGTGACCCCTATGTCAAAAAACTCTGTATGGAGAAGAGGAGAAATATGAGACAATGGTCACTTGT 189

Qy  178 AGCAAAATCCTCAGAAGGTGCCTAGTACTAAGCCATTGTGTTACTTAAGCCATGTGTTTGATGAGAGATGACT 237
        || |||| | ||||||  | ||| | | |||| |||| ||||   ||||| | | |||| ||||| | ||||
Db  178 AGCAAAATCCTCAGAAGGTGCCTAGTACTAAGCCATTGTGTTACTTAAGCCATGTGTTTGATGAGAGATGACT 237
```

RESULT 11

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 435
; CLASSIFICATION:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
Db  190 TTCACAATCCTAAGCAACAGTGCTTATGCATGAGGAGGATGTA---------- 229
        |||  |  || ||| || ||| |||  || | |||| |||| ||
Qy  238 AAAACAGGAGGCTGAAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCACGTTTGCTTGAA 297
        |||||  |  |  | ||  |||||||  |||| | ||| ||||||||
Db  230 ------------------ATGCTAAAACTCTCGCCAACAGAATTTGCTTGCT 261
Qy  298 GAAGAGATAAATGGATAACTAAATTAGAGATTAGAAGAAAATTAAGGATATGGAGTCACACAT 357
        ||||||  ||| ||| || ||||||  || ||| | |||||| ||||| | |||| ||||| 
Db  262 TGA---AGAAGAGTTCCTCAAGCAGTAATTAACGTTTGATTAGGATGTCACACAT 318
Qy  358 AATAAAGT---TTCTACCTTTCTTTAAAAGTGTAGCTAAAGTTGTGTTTTTTAATTGGCTTT 413
        |||||| |   |||| |||||||||| ||||| ||||| |||| | |||  || |||||| 
Db  319 AATTAAATAAAGTGTTGCCTTTTCTTATTAGGGTAACTTTATGATGTTGTATTTGGTTGGTTGGTTGGTA 378
Qy  414 TAGTAGCCCTTTTTTACTTACACTTAAAATAAGTAAGTGTGGCCACTTCAAATCCTTGGTGCAAAATCTTGGC 473
        ||||| || ||  ||||| ||| |||| || || |||||| || | |  | | || | |||||  ||||| 
Db  379 TAGTAGCCATTTGACACATCATCAAATAAATTAAAGTGTTGGACACATTGTGCAATCCTTGTGGTGTATCTTCT 438
Qy  474 ACTAAAGTTTATTTGTGTACTTTAATGAAAAATGACCTTCTATGGTCTTTGGTTAAAAAA 533
        ||  | |||||||| ||| |||   ||||| || || |||| | || |||||||||||||
Db  439 ATTAAG----TTTGTGTGTTTAAATGAAAAGACGTTTACGGTTCTAAAAAAAAAAAAA 494
Qy  534 AAAAAAAA 541
        ||||||||
Db  495 AAAAAAAA 502
```

```
;       IMMEDIATE SOURCE:
;         CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match      8.4%;  Score 45.4;  DB 1;  Length 7218;
Best Local Similarity 5.5%;  Pred. No. 0.0076;
Matches  19;  Conservative 184;  Mismatches 140;  Indels   0;  Gaps   0;

Qy      64 GAGGTGCAAGCTAGAGAAATGCAAAAACAGAAGCAACACATTCCTCGGAATATGCATTACC 123
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1367 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1308

Qy     124 AAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTACTAGGTGTCATTGTAGCCAAA 183
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1307 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1248

Qy     184 ATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATCAGAAGAGATGACTAAACA 243
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1247 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1188

Qy     244 GGAGCTGAAAATTTGGCTGAGGAAGCAAAAAACTTTGGCTGAGCTTTGCTTGAAGAGAG 303
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1187 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1128

Qy     304 ATAAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGATGCAGTATCACACATAATAAA 363
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1127 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1068

Qy     364 GTTTCTCACCTTTCTTAAAAGTGTAGCTAAATGTTGTGTTTTTAAT 406
              :::::::::::::::::::::::::::::::::::::::::
Db    1067 RATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGAATTAATT 1025
```

RESULT 12
US-09-601-198-63/c
; Sequence 63, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Elison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-63

Query Match   7.8%; Score 42; DB 4; Length 1692;
Best Local Similarity 48.7%; Pred. No. 0.039;
Matches 114; Conservative   0; Mismatches 120; Indels   0; Gaps   0;

```
Qy     303 GATAAATGGATAACTAATTAGAGATTAGAGAAAATTAAGGATGCAGTATCACACATAATAA 362
Db     871 GATAATAACTATTTATTAAAAGGTTATTATTTAAAAATCAATTATAAAATTGTCTTAATTCTTTAA 812

Qy     363 AGTTTCTCACCTTTCTTAAAAGTGTTAGCTATAATGTTGTGTTTTTAAGTTAGTAGCCT 422
Db     811 CAGCTAATTGGTTATTATTATAAAAATCATTAAAAATCATGGTTAAATTGTCTTAATTCTTTAA 752

Qy     423 TTTATTACACTTTAAAATAAGTGTGGCACTTCAATCCTTTGTTGCAATCTTTGCACTAAGTTT 482
Db     423
```

US-09-601-198-63

```
Db     751 TTTTGTCATCAGTCAAAAAGGTGATTTCTTGCACTTTTGTGAAGGCTATTTTTTATTAGCTT 692
Qy     483 ATTTGTAGTACTTTTTAATGAAAATGACCTTCTAATGGTTCTTATGGTCTTTGGTTAAAAAAAA 536
Db     691 CTTCAATAATTTTGATTGAATTTAAAAACTACAATGATATTTTTTAATAAATGAA 638
```

RESULT 13
US-09-801-861-3
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match   7.0%; Score 38; DB 4; Length 53332;
Best Local Similarity 51.8%; Pred. No. 1.4;
Matches 86; Conservative   0; Mismatches 80; Indels   0; Gaps   0;

```
Qy     332 GAAATTAAGGATGCAGTATCACACATAAATAAAAGTTTCTCACCTTTCTTAAAAGTGTTAGCTA 391
Db   31303 GCAATCACGTATATATTTATTTTTTATATTTATTTAATATTTTTATAIATAAGTGTGGCACT 31362

Qy     392 ATGTTGTGTTTTTAATTGGCTTTTAGTAGCCTTTAATTACACTTTAAATAAGTGTGGCACT 451
Db   31363 GTTATATATTTTATATATTTTTTATATATTTTGTATATATATTTATATAT 31422

Qy     452 TCAATCCTTTGTTGCAATCTTGCACTAAGTTTAATTTTGTGTGTAGCTTTTA 497
Db   31423 TTATATATTTTTATATACATTATATTTTTTATATATTTATATTTTTA 31468
```

RESULT 14
US-08-472-028A-3
; Sequence 3, Application US/08472028A
; Patent No. 5673373
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Volrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,028A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1748/CIP
TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
 LENGTH: 1738 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
 NAME/KEY: CDS
 LOCATION: 70..1596
 OTHER INFORMATION:
US-08-472-028A-3   OTHER INFORMATION: /note= "Arabidopsis protox-2 cDNA;
                                    sequence from pWDC-1"

Query Match   6.9%; Score 37.2; DB 1; Length 1738;
Best Local Similarity 54.3%; Pred. No. 0.76;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy   253 ATTTTGGCTGAGGAAGCAAAAAACTTTGGCTCCAGCTTTGCTTGAGGAAGAGATAAATGGAT 312
         ||||||||| |||||||||||||||||||| |||||||||| || |||||||||| |||||
Db   259 ATTTTGGGATGAAGGAGGCAAACACCATGACTGAGGCTGAGCCAGAAGTTGGGAGTTTACTT 318

Qy   313 AACTAATTAGAGATTAGAGAAGAAATTAAGGATGCAGTATCACACACATAATAAAAGTTTCTACC 372
         |||||||||||||| |||||||||||||||||| |||||||| |||||||||||||||||||||
Db   319 GATGATCTTGGGCTTCGTGAGAAACAACAATTTCCAATTTCACAGAAAAAGCCGGTATATT 378

Qy   373 TTTCTTAAAAGTGTAGCT 390
         |||||||| |||| |||
Db   379 GTGCGGAATGGTGTACCT 396

RESULT 15
US-08-808-931-3
Sequence 3, Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
 APPLICANT: Volrath, Sandra
 APPLICANT: Johnson, Marie
 APPLICANT: Potter, Sharon
 APPLICANT: Ward, Eric
 APPLICANT: Heifetz, Peter
 TITLE OF INVENTION: DNA Molecules Encoding Plant
 TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
 TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5939602artis Corporation
 STREET: 520 White Plains Road, P.O. Box 2005
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591-9005
COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/808,931
 FILING DATE:
PRIOR APPLICATION DATA:
 CLASSIFICATION: 435
 APPLICATION NUMBER: US 60/012,705
 FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/013,612
 FILING DATE: 28-FEB-1996

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/020,003
 FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919) 541-8689
 TELEFAX: (919) 541-8587
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
 LENGTH: 1738 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
 ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
 CLONE: pWDC-1 (NRRL B-21237)
FEATURE:
 NAME/KEY: CDS
 LOCATION: 70..1596
US-08-808-931-3   OTHER INFORMATION: /product= "Arabidopsis protox-2"

Query Match   6.9%; Score 37.2; DB 2; Length 1738;
Best Local Similarity 54.3%; Pred. No. 0.76;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy   253 ATTTTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTTGCTTGAGGAAGAGATAAATGGAT 312
         ||||||||| |||||||||||||||||||| |||||||||| || |||||||||| |||||
Db   259 ATTTTGGGATGAAGGAGGCAAACACCATGACTGAGGCTGAGCCAGAAGTTGGGAGTTTACTT 318

Qy   313 AACTAATTAGAGATTAGAGAAGAAATTAAGGATGCAGTATCACACACATAATAAAAGTTTCTACC 372
         |||||||||||||| |||||||||||||||||| |||||||| |||||||||||||||||||||
Db   319 GATGATCTTGGGCTTCGTGAGAAACAACAATTTCCAATTTCACAGAAAAAGCCGGTATATT 378

Qy   373 TTTCTTAAAAGTGTAGCT 390
         |||||||| |||| |||
Db   379 GTGCGGAATGGTGTACCT 396

Search completed: January 28, 2004, 09:56:13
Job time : 63 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 08:44:46 ; Search time 1946 Seconds
                 (without alignments)
                 6756.794 Million cell updates/sec

Title: US-10-072-809B-17
Perfect score: 541
Sequence: 1 atggctcgggctccttggtgctt.............tggttaaaaaaaaaaaaaa 541
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152228056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
        1: em_estba:*
        2: em_esthum:*
        3: em_estli:*
        4: em_estmu:*
        5: em_estov:*
        6: em_estpl:*
        7: em_estro:*
        8: em_htc:*
        9: gb_est1:*
       10: gb_est2:*
       11: gb_htc:*
       12: gb_est3:*
       13: gb_est4:*
       14: gb_est5:*
       15: em_estrun:*
       16: em_estcom:*
       17: em_gss_hum:*
       18: em_gss_inv:*
       19: em_gss_pln:*
       20: em_gss_vrt:*
       21: em_gss_fun:*
       22: em_gss_mam:*
       23: em_gss_mus:*
       24: em_gss_pro:*
       25: em_gss_rod:*
       26: em_gss_phg:*
       27: em_gss_vrl:*
       28: em_gss1:*
       29: gb_gss2:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229.2	42.4	316	13	BU494528	DD-R7-01
2	200.4	37.0	570	12	BI929149	EST549038
3	198.8	36.7	570	12	BI927088	EST546977
4	198.8	36.7	570	12	BI928623	EST548812

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

No.	Score	%	Length	DB	ID	Description
5	196.4	36.3	533	12	BI927975	EST547852
6	194	35.9	541	12	BI929019	EST548908
7	194	35.9	541	12	BI929156	EST549045
8	189.8	35.1	522	9	AI485044	EST243324
9	189.8	35.1	522	9	AI485044	EST243871
10	189.8	35.1	522	9	AI487398	EST245720
11	189.8	35.1	551	9	AI489818	EST248157
12	189.8	35.1	562	9	AW217379	EST296063
13	189.8	35.1	564	9	AI897456	EST266899
14	189.8	35.1	712	9	AI898009	EST267452
15	188.2	34.8	500	9	BI931665	EST515554
16	188.2	34.8	594	10	AW623541	EST321486
17	188.6	34.5	562	12	AW929929	EST354199
18	186.6	34.3	672	12	AI486656	EST354977
19	185.8	34.3	567	9	BM064987	KS01073F0
20	183.8	34.0	561	12	AI485550	EST243871
21	183	33.8	561	12	BI927588	EST547489
22	182.8	33.8	477	9	AI897392	EST266835
23	182	33.6	419	9	AW399651	EST310151
24c	181.8	33.6	480	9	AI490243	EST248569
25	179.4	33.2	563	9	AI483612	EST249462
26c	176.8	32.7	574	12	BM067544	KS0807B0
27	170	31.4	378	9	AI484887	EST243150
28	170	31.4	586	9	AI487824	EST246146
29c	169.2	31.3	435	10	AW929939	EST354210
30c	167	30.9	481	10	BG628107	CC-esf1cL
31	166.4	30.9	513	12	BM063138	KS01052A0
32	167	30.9	511	12	BM068404	KS08023B0
33c	163.6	30.2	491	12	BM067238	KS08001C0
34	162.6	30.1	573	14	CA523446	KS12022C1
35	160.8	29.7	403	12	BI929011	EST548909
36c	160	29.7	403	12	BM061668	KS01033C1
37	158.2	29.6	500	12	AI487859	EST241181
38	158.4	29.2	505	9	AI898075	EST267518
39	158	28.9	421	9	AI898075	EST241181
40c	154.8	28.6	536	12	BM063253	KS01053C0
41c	152.2	28.1	589	14	CA516706	KS09062G0
42	148.6	27.5	333	14	AW623112	EST321057
43	145.8	27.0	476	10	BG629274	CC-esf1cL
44	137.8	25.5	537	12	BM064540	KS01068D0
45	134.6	24.9	532	12	BM063609	KS01057D1
	133	24.6	571	14	CA515866	KS09048E1
	132.2	24.4	603	9	AI483999	EST249870

ALIGNMENTS

RESULT 1
BU494528
BU494528

```
LOCUS       BU494528                316 bp    mRNA    linear   EST 09-JAN-2003
DEFINITION  DD-R7-01 DD-R7-derived M.sexta/T.notatus-responsive library
            Nicotiana attenuata cDNA clone cv84.4 3' similar to Solanaceae mRNA
            for flower-specific thionin (Z11748.1), mRNA sequence.
ACCESSION   BU494528
VERSION     BU494528.1  GI:27552699
KEYWORDS    EST.
SOURCE      Nicotiana attenuata
  ORGANISM  Nicotiana attenuata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 316)
  AUTHORS   Voelckel,C. and Baldwin,I.T.
  TITLE     Detecting herbivore-specific transcriptional responses in plants
            with multiple DDRT-PCR and subtractive library procedures
  JOURNAL   Unpublished
COMMENT     Contact: Gase K
            Department of Molecular Ecology
            Max Planck Institute for Chemical Ecology
            Winzerlaer Str. 10, Beutenberg Campus, 07745 Jena, Germany
            Tel: 49/3641/571121
            Fax: 49/3641/571102
            Email: gase@ice.mpg.de
            Seq primer: M13 Forward (-20)
```

FEATURES
  source   Location/Qualifiers
                  1..316
                  /organism="Nicotiana attenuata"
                  /mol_type="mRNA"
                  /clone="cv84.4"
                  /db_xref="taxon:49451"
                  /tissue_type="shoot"
                  /dev_stage="rosette"
                  /clone_lib="DD-R7-derived M.sexta/T.notatus-responsive
                  library"
                  /lab_host="E.coli"
                  /note="Vector: pCR 2.1-TOPO; Site_1: EcoR I; Site_2: EcoR
                  I; DDRT-PCR with arbitrary primer_R7 (Nucleic Acid
                  Research 21, 4272-4280, 1993)"

POLYA=Yes.

BASE COUNT   97 a   44 c   66 g   109 t
ORIGIN

Query Match   42.4%;   Score 229.2; DB 13; Length 316;
Best Local Similarity   91.4%;   Pred. No. 2.4e-31;
Matches 265; Conservative   0; Mismatches   23; Indels   2; Gaps   2;

Qy   239 AAACAGGAGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAG 298
                ||||||||||||||||| || |||||||||||||||||||||||||||||||||||||
Db   4 ATACAGAGAGCTGAAACTTTAGCCGAGGAGAGCAAAACACTTTGGCTGCAGCTTTGCTTGAAG 63

Qy   299 AAGAGGATAATGGATAACTAATTAGAGGATTAGAAGAAAATTAAGGATGGCAGTATCACACATA 358
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   64 AAGAGATGATGGATAACTAATTGGGGATTAGAAGAAAATTAAGGATGAAAGTGTCACACATA 123

Qy   359 ATAAAAGTTTCTACCTTTCTTAAAAGAGTTAGCTAAGTGTAGCTAAAGTTGGTTTTAATTGG 418
                ||||||| ||||||||||||||||||||||||| || || |||||||||||||||
Db   124 ATAAAAG-TGCTGCCTTTTCTTATAGTGTAGCTAAATGTTGTGTTTCTTTATTGCGACTTGA 182

Qy   419 GCCTTTTAATACACTTTAAATAAGTGTGTGGCCACTTCAAATCCTTTGTGCAAATCTTGCTAA 478
                ||||||||||||||||||||||||||||||||||||| ||||||| |||||||
Db   183 GCCGTTGTTACACTTTAAATAAGTGTGGCCACGTCAATCCTTTGTGCAAATCTTGTACTAA 242

Qy   479 GTTTATTTGTGTGTACTTTTTAATGAAAAAATGAACCTTCTATGGTTCTTGGTTAA 528
                ||||| | ||| || ||||||||||||||||||||||||||||||||||||||
Db   243 GTTTATTTGTGTGTACTTTTT-ATGAAAATGAACCTTCTATGGTTCTTAGGTTAA 291

RESULT 2
BI929149
LOCUS   BI929149
DEFINITION   EST549038 tomato flower, 3'-8 mm buds Lycopersicon esculentum cDNA
                clone cTOB27L16 5' end, mRNA sequence.
ACCESSION   BI929149
VERSION   BI929149.1 GI:16241157
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
  ORGANISM   Lycopersicon esculentum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 570)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
                Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
                Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE   Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished
  COMMENT   Contact: CUGI
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson, SC 29634, USA
                Email: http://www.genome.clemson.edu/orders/index.html
                This clone is available through the Clemson University Genomics
                Institute
                Seq primer: T3.

FEATURES
  source   Location/Qualifiers
                  1..570
                  /organism="Lycopersicon esculentum"
                  /mol_type="mRNA"
                  /cultivar="TA496"
                  /db_xref="taxon:4081"
                  /clone="cTOB27L16"
                  /tissue_type="flower"
                  /dev_stage="3-8mm buds"
                  /clone_lib="tomato flower, 3 - 8 mm buds"
                  /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                  XhoI; supplier: Cornell University; sequencing; The
                  Institute for Genomic Research. Flower buds and flowers
                  were taken from greenhouse plants (4-8 wks old, TA496).
                  They were immediately frozen in liquid nitrogen and then
                  size-separated while remaining frozen."

BASE COUNT   179 a   81 c   119 g   191 t
ORIGIN

Query Match   37.0%;   Score 200.4; DB 12; Length 570;
Best Local Similarity   68.5%;   Pred. No. 3e-26;
Matches 370; Conservative   0; Mismatches 151; Indels   19; Gaps   6;

Qy   1 ATGGCTCGCTCCTTGTGCTTCATCGGCATTTGCTATCTTGGCAAGGATGCTTGTTGTTGCC 60
                ||||| || || || | |||||| || |||| |||||| ||| ||||||||||||||
Db   34 ATGGCTCCGTTCCATTTGCTTCATCATGGCCACTTATGGTTCTTGGCAATCGTGGCTCTTTGTTCC 93

Qy   61 TATGAGGTGCAAGCTAGAGA---ATGCAAAACACGAAAGCAACACATTTCCTCGGAATATGC 117
                |||||||||||||||||| |||||| || || || ||||| ||||||| ||||
Db   94 TCTGAGGTGCAAGCTCAGCAGATGTGCAAATCAACAACGCAAAAAACCTTCAAGGGAATATGC 153

Qy   118 ATTACCAAACACCATGCAGAAAAAGCTTGTATCAGTGAGAAAATTTACTGATGTCATTGT 177
                |||||||||||| |||| |||||||||||||| |||||| ||| ||| || |||||||
Db   154 TTTACCGATTCATCGTGTAGAAAGCCTTGTGTCACAGAGGAGTTACAGGTGGACATTGT 213

Qy   178 AGCAAAATCCTCAGAAGGGTGCCTATGTAGTAACCATGGTAGCTAAGCCATGGATGAAGAGGATGACT 237
                || || | ||| ||||| || ||| || |||| ||| ||| | || ||||| || |||
Db   214 AGCAAAACTCCAAAGAAAGGTGCCCATAGTACTAACAAGGTTGTAATTTGAAAGAAGTGATCAAAT 273

Qy   238 AAAACAGGAGCTGAAAATTTTGGCTGAGGAAGCAAAAACACTTTGGCTGCAGCTTTGCTTGAA 297
                ||| ||||| || |||||||| ||| |||| ||||| ||| || |||| || ||||||||
Db   274 GAAGTTAAAAC---AACTTTGGGTTGGGGAAGCAAAAAACTTAAGTGAAAACTGTGGTTGAA 330

Qy   298 GAAGAGATAATGGATAACTAATTA------GAGATTAGAAGAAAATTAAGGATGGCAGTATC 351
                ||||||||| |||||| ||||| ||| ||| ||||||||||| || |||||||
Db   331 GAAGAGATTGTGATGGAGTAATTAATTAATAAGTGAGATAATTAAGAATTTGAGTGTCAAAATC 390

Qy   352 ACACATAATAAAAGTTTCTACCTTTCTTAAAAAGAGTGTAGCTAAAGTTGGTTTTTAATTGGCT 411
                || |||| |||||| || |||| ||||| |||| ||| || || ||||| ||||| |||
Db   391 AAAATTAATAAAAGTTAATTGTCTTTTACTTAGTAAAAGGGTAGCTTCTAAATGTCGTGTTATTGGGC 450

Qy   412 TTTAGTAGCCTTTATTACACTTTAAATAAGGTGCGCACTTCAAT---CCTTGTGCAAT 468
                |||| || |||||||| |||| |||||| ||| || |||| ||||| |||
Db   451 TTTAGTAGCCANTTGATACA--TTAAATAAGTTTTGACATGCATAAATCCTTTTTGGATT 508

Qy   469 CTTGCACTAAGTTATTTGTGTA--CTTTTAATGAAAAATGACCTTCTATGGTTCTTGGTTT 526
                |||| ||| ||||||| ||| |||| || ||||| |||||| |||||||| |||
Db   509 CTTGTATCAACTTAATCAGTGTAAGTTTTTATGAAAAATGACTATGGTGTCTTTAATT 568

RESULT 3
BI927088
LOCUS   BI927088
DEFINITION   EST546977 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
                clone cTOB20I01 5' end, mRNA sequence.
ACCESSION   BI927088
VERSION   BI927088.1 GI:16236262
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
  ORGANISM   Lycopersicon esculentum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 570)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
                Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,

RESULT 4
BI928623

```
TITLE       Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL     Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..570
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOB2101"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT    180 a     81 c    119 g    190 t
ORIGIN

Query Match     36.7%;  Score 198.8;  DB 12;  Length 570;
Best Local Similarity   68.3%;  Pred. No. 5.8e-26;
Matches 369;  Conservative   0;  Mismatches 152;  Indels  19;  Gaps   6;

Qy    1 ATGGCTGCCTCCTTGTGTGCTTCATGGCCATTTGCTATCTTGGCAAGGATGGCTGCTTTTGTTGCC 60
        ||||||   ||  ||| || ||  |||  ||       | |  ||||   ||||||||||| ||
Db   34 ATGGCTCGTTCCATTGTTCTCATGGCCACTATGGTCTTGGCAAAAGATGGCTGCTTTTGTTCC 93

Qy   61 TATGAGGTGCAAGCTAGAGA---ATGCAAAACAGAAAGCAACACACATTTCCTGGAATATGC 117
        ||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db   94 TATGAGGTGCAAGCTCAGCAGATGTGCAAAATCAACAACCAAAACCTTCAAGGGATTATGC 153

Qy  118 ATTACCAAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
        ||||||||||||||||||||||||||||||||       ||||||||||  |||||||||
Db  154 TTTACCGATTCATCCGTGTAGAAAGGCTTGTGTCACAGAGGATGTTGGACATTGT 213

Qy  178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTTGATGAAGATGACT 237
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  214 AGCAAACTCCAAAGAAAGTGCCTATGTACTAAGGTTGTGTATTTGAAAAGAATCAAAT 273

Qy  238 AAAACAGGAGCTGAAATTTGGCTGAGGAAGCAACACATTTGGCTGCTTGCTTGAA 297
        |||||||||||||   |||||||||||||||||||||||||||||||||||||
Db  274 GAAGTTAAAAC---AACTTGGTTGGGGAAGCAAAACTCTAAGTGAAACTGTGCTTGAA 330

Qy  298 GAAGAGAATAATGGATAACTAAATTA------GAGATTAGAGAAAATTAAGGATGTCAGTTGT 351
        ||||||||||||||||||||||||       |||||||||||||||||||||||||||||
Db  331 GAAGAGATTGTGATGGAGTAATTAATAAGGAATTTGAGGTGTCAAAAATC 390

Qy  352 ACACATAATAAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTGTGTTTAATTGGC 411
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  391 AAAATTAATAAAGTATTGTCTTTTAGTAAAAGGGTAGCTTCTAAATGTCGTGTTATTGGGC 450

Qy  412 TTTAGTAGCCTTTATTACACTTTAAATAAGTGTGGCACTTCAAT---CCTTTGTGCAAT 468
        ||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db  451 TTTAGTAGCCATTTGATACA---TTAAATAAGTTTGACATGTCATTAATCCTTTGATT 508

Qy  469 CTTGCACTAAGTTTATTTATTTGTGTA---CTTTTAATGAAAAATGACCTTCTATGGTCTTTTGGTT 526
        ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db  509 CTTGTATCAACTTAATCAGTGTAAGTTTAATGAAAAATGATCGACTATGGTCTTTAATT 568
```

```
LOCUS       BI928623                570 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST548512 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
            clone cTOB26K7 5' end, mRNA sequence.
ACCESSION   BI928623
VERSION     BI928623.1  GI:16240393
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..570
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOB26K7"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT    180 a     81 c    119 g    190 t
ORIGIN

Query Match     36.7%;  Score 198.8;  DB 12;  Length 570;
Best Local Similarity   68.3%;  Pred. No. 5.8e-26;
Matches 369;  Conservative   0;  Mismatches 152;  Indels  19;  Gaps   6;

Qy    1 ATGGCTGCCTCCTTGTGTGCTTCATGGCCATTTGCTATCTTGGCAAGGATGGCTGCTTTTGTTGCC 60
        ||||||   ||  ||| || ||  |||  ||       | |  ||||   ||||||||||| ||
Db   34 ATGGCTCGTTCCATTGTTCTCATGGCCACTATGGTCTTGGCAAAAGATGGCTGCTTTTGTTCC 93

Qy   61 TATGAGGTGCAAGCTAGAGA---ATGCAAAACAGAAAGCAACACACATTTCCTGGAATATGC 117
        ||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db   94 TATGAGGTGCAAGCTCAGCAGATGTGCAAAATCAACAACCAAAACCTTCAAGGGATTATGC 153

Qy  118 ATTACCAAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
        ||||||||||||||||||||||||||||||||       ||||||||||  |||||||||
Db  154 TTTACCGATTCATCCGTGTAGAAAGGCTTGTGTCACAGAGGATGTTGGACATTGT 213

Qy  178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTTGATGAAGATGACT 237
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  214 AGCAAACTCCAAAGAAAGTGCCTATGTACTAAGGTTGTGTATTTGAAAAGAATCAAAT 273

Qy  238 AAAACAGGAGCTGAAATTTGGCTGAGGAAGCAACACATTTGGCTGCTTGCTTGAA 297
        |||||||||||||   |||||||||||||||||||||||||||||||||||||
Db  274 GAAGTTAAAAC---AACTTGGTTGGGGAAGCAAAACTCTAAGTGAAACTGTGCTTGAA 330

Qy  298 GAAGAGAATAATGGATAACTAAATTA------GAGATTAGAGAAAATTAAGGATGTCAGTTGT 351
        ||||||||||||||||||||||||       |||||||||||||||||||||||||||||
Db  331 GAAGAGATTGTGATGGAGTAATTAATAAGGAATTTGAGGTGTCAAAAATC 390

Qy  352 ACACATAATAAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTGTGTTTAATTGGCT 411
```

RESULT 5

```
BI927975
LOCUS       BI927975                 533 bp    mRNA    linear    EST 18-OCT-2001
DEFINITION  EST547B52 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
            clone cTOB24E21 5' end, mRNA sequence.
ACCESSION   BI927975
VERSION     BI927975.1  GI:16238542
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum
  ORGANISM  Lycopersicon esculentum (tomato)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 533)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..533
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /clone="cTOB24E21"
                     /db_xref="taxon:4081"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      168 a     113 g     72 c     180 t
ORIGIN
```

```
Qy    391 AAAATTAATAAAAGTATTGTCTTTAGTAGAAAAGGGTAGCTTCTAATGTCGTGTGTATTGGGC 450
          |||||||||||||||||||||||||||||| |||||||||  ||| ||| ||||||  |||||
Db    412 TTTAGTTAGCCTTTATTACACTTTAAATAAGTAGTGGCACTTCAAT---CCTTTGTGTCAAT 468

Qy    451 TTTAGTTAGCCATTTGATACA--TTAAATAAGTTTTGACATGTCATTAATCCTTTGGATT 508
          ||||||||||||||||||||||  ||||| ||||||| |||||||||||||||||||
Db    451 TTTAGTTAGCCATTTGATACA--TTAAATAAGTTTTGACATGTCATTAATCCTTTGGATT 508

Qy    469 CTTGACTAAGTTATTTGTGTA--CTTTTAATGAAAATGACCTTCTATGGTCTTTTGGTT 526
          |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    509 CTTGTATCACTTAATCACAGGTAAGTTTAATGAAAATGATCGACTATGGTCTTTTAATT 568
```

RESULT 6

```
BI929019
LOCUS       BI929019                 541 bp    mRNA    linear    EST 18-OCT-2001
DEFINITION  EST548908 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
            clone cTOB27M14 5' end, mRNA sequence.
ACCESSION   BI929019
VERSION     BI929019.1  GI:16241728
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum
  ORGANISM  Lycopersicon esculentum (tomato)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..541
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /clone="cTOB27M14"
                     /db_xref="taxon:4081"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      170 a     115 g     78 c     178 t
ORIGIN
```

```
Qy    182 AAATCCTCAGAAGGTGCCTAGTGTACTAAGGCCATGTGTGTTCTTTGATGAGAAGATGACTAAAA 241
          ||||||||||||||||||||||||||||||||||||||  ||||| |||||||| ||||||||||
Db    181 AACTCCAAAGAAAGTGGCCTATGTTACTAAGGTTTGGTGTATTTGAAAAAGATTCAAATGAAG 240

Qy    242 CAGGAGGCTGAAATTTGGCTGAGGAAGGCCAAAAACTTTGGCGTCGACGTTTTGCTTGTGAGGAG 301
          |||||||||||||||||||||||||||||||| ||||||||||||||||||  ||| ||||||||
Db    241 TTAAAAC---AACTTTGGTTGGGGAGCAAGCAAAAACTCTAAGTGAAACTGTGCGTCCTTGAAGAAG 297

Qy    302 AGATAATGGATAACTAATTA------------GAGATTAGAGAAAATTAAGGATGCAGTATCACAC 355
          |||||||||||||||| ||||         ||||| |||||||||  |||||| ||||||||||
Db    241 TTAAAAC---AACTTTGGTTGGGGAGCAAGCAAAAACTCTAAGTGAAACTGTGCGTCCTTGAAGAAG 297

Qy    356 ATAATAAAGTTTCTACACCTTTCTTAAAAGTGAGACTAATGTGTGTTTTAATTGGCTTTTA 415
          |||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||
Db    358 TTAATAAAGTATTGTCTTTAGTAGAAAGGGTAGCTTCTAAATGTCGTGTGTTATTGGGCTTTA 417

Qy    416 GTAGCCTTTATTACACTTTAAATAAGTGGCACTTCAAT---CCTTTGTGCAATCTTG 472
          |||||||||||||| |||||||||||||| ||||||||   |||||||||||||||
Db    418 GTAGCCATTTGATACA--TTAAATAAGTTTTGACATGTCATTAATCCTTTGGATTCTTG 475

Qy    473 CACTAAGTTATTTGTGTA--CTTTTAATGAAAATGACCTTCTATGGTCTTCTTGGTT 526
          ||||| ||||||||||||  ||||||||||||| |||| ||||||||||  |||||
Db    476 TATCACTTAATCACAGGTAAGTTTAATGAAAATGATCGACTATGGTCTTTTAATT 531
```

BASE COUNT
ORIGIN

```
Qy    5 CTCGTCCTTGTGTGCTTCATGGCATTGCTATCTTGGCAAGGATGCTCTTTGTTGTTGCCTATG 64
          |||||||| ||||| |||| ||||| |||||||||||||||||||||||||||||||||||
Db    1 CTCGTTCCATTTGTTCATCGACTTATGGTCTTGGCCAATGGTCCTCTTTGTTCCTCTG 60

Qy   65 AGGTGCAAGCTAGAGA---ATGCAAAACAGAAAAGCAACACATTTCCTGGAAATATGCATTA 121
          |||||||||||||||    |||| ||||||||||||||||||||||||||||||||
Db   61 AGGTGCAAGCTCAGCAGATGCAAATCAACAAACCTTCAAGGGATTATGCCTTTA 120

Qy  122 CCAAACCACCATGCAGAGAAAAAAGCTTGTATCGATGGTCATTGGTTAGCA 181
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CCGATTCATCGTGTAGAGAAAGGCTTGTGTCACAGAGGAGTTTACAGGTGGACATTGTGAGCA 180
```

```
Query Match      36.3%;  Best Local Similarity  68.3%;  Conservative  0; Mismatches 151; Indels 19; Gaps 6;
Matches 366; Conservative  0; Mismatches 151; Indels 19; Gaps 6;
Best Local Similarity  68.3%;  Score 196.4;  DB 12;  Length 533;

Query Match   35.9%;   Score 194;  DB 12;  Length 541;
```

RESULT 7

```
Best Local Similarity: 68.2%;   Pred. No. 4.3e-25;
Matches 348; Conservative  0; Mismatches 145;  Indels  17; Gaps  5;

Qy      1 ATGGCTCGGTTCCATTGTGCTTCATGGCAATTGCTATCTTGGCAAGGATGCTTCTTGTTGCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGGCTCGGTTCCATTGTGCTTCATGGCACTTATGGTGTCTTGGCAAATGCTTCTTGTTCC 93

Qy     61 TATGAGGTTGCAAGCTAGAGA---ATGCAAAACAGAAAGCAACACATTTCCTGGAATATGC 117
          |||||||||||||||||||       |||||||||||||||||||||||||||||||||
Db     94 TCTGAGGTTGCAAGCTCAGCAGATGTGCAAAATCAACAAGCCAAAACCTTCAAGGGAATTATGC 153

Qy    118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAAATTACTGATGGTCATTGT 177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TTTACCGATTCATCGTGTAGAAAGGCTTGTCACGAGGAGTTACAGGGTGGACATTGT 213

Qy    178 AGCAAAATCCTCAGAAAGGTGGCCTATTGACTAAGGCCATGTGTGTTGGATGAAGATGACT 237
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    214 AGCAAACTCCAAAGAAGAGAGTTGCCTATTGACTAAGGCCATGTGTGTATTTGTAATAAAAT 273

Qy    238 AAAACAGGAGCTGAAAATTTGGCTGAGGAAGGCGAAAAAACTTGGGCTTGCTTGAA 297
          ||||||||||||||||||||||||||||||||||||||||||
Db    274 GAAGTTAAAAC---AACTTGGTTGGGGAAGCAAAAAACTCTAAGGTGAAACTGCTTGCTTGAA 330

Qy    298 GAAGAGATAAATGGATAACTAATTA------GAGATTAGAAGAAAATTAAGGATGCAGTATC 351
          ||||||||||||||||||||||||||       ||||||||||||||||||||||||
Db    331 GAAGAGATTGTGATGGAGTAATTAAAATAAGTGAGATAATTAATTAAGAATTTGAGTGTCAAAATC 390

Qy    352 ACACATAATAAAAGTTTCTACCTTTCTTAAAAGTTTGAGCTAATGTTGTTTTTAATTGGCT 411
          |||||||||||||||||||||||||||||||||||||||||||||
Db    352 ACACATAATAAAAGTTTCTACCTTTCTTAAAAGTTTGAGCTAGTTGTTGTTTTAATTGGGC 411

Qy    412 TTTAGTAGCCTTTTATTACACTTTAAAATAAGTGTGGGCACTTCAAT---CCTTTGTGCAAT 468
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 TTTAGTAGCCATTTGATACA--TTAAATAAGTTTTGACAATGTCATTAATCCTTTTTGGATT 508

Qy    469 CTTGCACTAAGTTTATTTGTGTACTTTTTAA 498
          ||||||||||||||||||||||||||||||
Db    509 CTTGTATCAACTTAATCAGTGTAAGTTTTA 538
```

LOCUS   BI929156   541 bp   mRNA   linear   EST 18-OCT-2001
DEFINITION BI929156 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
            clone cTOB27N10 5', mRNA sequence.
ACCESSION   BI929156
VERSION   BI929156.1 GI:16242173
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
  ORGANISM Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 541)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE   Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished
  COMMENT   Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES   Location/Qualifiers
     source   1..541
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOB27N10"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /note="tomato flower, 3 - 8 mm buds"
                     /clone_lib="tomato flower"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT   170 a   78 c   115 g   178 t
ORIGIN

RESULT 8

```
Query Match      35.9%;  Score 194;  DB 12;  Length 541;
Best Local Similarity  68.2%;   Pred. No. 4.3e-25;
Matches 348; Conservative  0; Mismatches 145;  Indels  17; Gaps  5;

Qy      1 ATGGCTCGGTTCCATTGTGCTTCATGGCAATTGCTATCTTGGCAAGGATGCTTCTTGTTGCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGGCTCGGTTCCATTGTGCTTCATGGCACTTATGGTGTCTTGGCAAATGCTTCTTGTTCC 93

Qy     61 TATGAGGTTGCAAGCTAGAGA---ATGCAAAACAGAAAGCAACACATTTCCTGGAATATGC 117
          |||||||||||||||||||       |||||||||||||||||||||||||||||||||
Db     94 TCTGAGGTTGCAAGCTCAGCAGATGTGCAAAATCAACAAGCCAAAACCTTCAAGGGAATTATGC 153

Qy    118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAAATTACTGATGGTCATTGT 177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TTTACCGATTCATCGTGTAGAAAGGCTTGTCACGAGGAGTTACAGGGTGGACATTGT 213

Qy    178 AGCAAAATCCTCAGAAAGGTGGCCTATTGACTAAGGCCATGTGTGTTGGATGAAGATGACT 237
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    214 AGCAAACTCCAAAGAAGAGAGTTGCCTATTGACTAAGGCCATGTGTGTATTTGTAATAAAAT 273

Qy    238 AAAACAGGAGCTGAAAATTTGGCTGAGGAAGGCGAAAAAACTTGGGCTTGCTTGAA 297
          ||||||||||||||||||||||||||||||||||||||||||
Db    274 GAAGTTAAAAC---AACTTGGTTGGGGAAGCAAAAAACTCTAAGGTGAAACTGCTTGCTTGAA 330

Qy    298 GAAGAGATAAATGGATAACTAATTA------GAGATTAGAAGAAAATTAAGGATGCAGTATC 351
          ||||||||||||||||||||||||||       ||||||||||||||||||||||||
Db    331 GAAGAGATTGTGATGGAGTAATTAAAATAAGTGAGATAATTAATTAAGAATTTGAGTGTCAAAATC 390

Qy    352 ACACATAATAAAAGTTTCTACCTTTCTTAAAAGTTTGAGCTAATGTTGTTTTTAATTGGCT 411
          |||||||||||||||||||||||||||||||||||||||||||||
Db    352 ACACATAATAAAAGTTTCTACCTTTCTTAAAAGTTTGAGCTAGTTGTTGTTTTAATTGGGC 411

Qy    412 TTTAGTAGCCTTTTATTACACTTTAAAATAAGTGTGGGCACTTCAAT---CCTTTGTGCAAT 468
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 TTTAGTAGCCATTTGATACA--TTAAATAAGTTTTGACAATGTCATTAATCCTTTTTGGATT 508

Qy    469 CTTGCACTAAGTTTATTTGTGTACTTTTTAA 498
          ||||||||||||||||||||||||||||||
Db    509 CTTGTATCAACTTAATCAGTGTAAGTTTTA 538
```

LOCUS   AI485044   522 bp   mRNA   linear   EST 18-MAY-2001
DEFINITION AI485044 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED2L8, mRNA sequence.
ACCESSION   AI485044
VERSION   AI485044.1 GI:4380415
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
  ORGANISM Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 522)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE   Generation of ESTs from tomato carpel tissue

RESULT 9

```
LOCUS       AI487398                522 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST245720 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED13B21, mRNA sequence.
ACCESSION   AI487398
VERSION     AI487398.1  GI:4382769
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1..522
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED13B21"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /lab_host="Xl1-Blue MRF'"
                     /clone_lib="tomato ovary, TAMU"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; cLED - Tomato Carpel EST Library. Oligodr-primed and
                     directionally cloned cDNA in vector lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
BASE COUNT      166 a      102 c      177 g      77 t
ORIGIN

Query Match             35.1%;  Score 189.8;  DB 9;  Length 522;
  Best Local Similarity   68.9%;  Pred. No. 2.5e-24;
  Matches  326;  Conservative    0;  Mismatches  127;  Indels   20;  Gaps    4;

Qy      1 ATGGGCTCGCTCCTTCCTTGTGCTTCAATGGCCATTTCTATCCTTGGCCAAGGATGCTCTTCTTTGTTGCC 60
          |||||||||| || || | |||||||| ||||||||
Db     27 ATGGGCTCGTTCCATTTTCTTCATGGCATT                                 TTTGTTACC 86

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTTCTGGATATGC 117
          |||||| ||| | ||||||||   ||||||||
Db     87 TATGAGGTTAGAAGCTTCATCAAATTTGCAAAGCACCAAACTTCCCAGGATTATGT 146

Qy    118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
          |||||||||||
Db    147 TTTATGGACTCATCATGTAGAAAATAITTGTATCAAAGAGAAAATTTACTGGTGGAACATTGT 206

Qy    178 AGCAAAAATCCTCAGAAGGTGCCTATGTACTTAAGCCACATGTGTGTTTTGATGAGAAGATGACT 237
          ||||||||
Db    207 AGCAAACTCCAAAGCAAGAAGTGTCTATGCACTTAAGCCACATGTGTGTTATT---TGACAAAAATCTCA 263

Qy    238 AAAACAGGAGCTGAAATTTGGGCTGAGGAAGCAAAAACTTTGGCCTTGCCTTGAA 297
          ||||||||
Db    264 AGTGAAGTTAAAGCAACACTTTGGGTGAGGAAGCAAAAACTCAAGTGAAGTTGTGCTTGAA 323

Qy    298 GAAGAGATAATGGATTAACTAATTAGAGAGATTAGAAGAAAATTAAGGAT 346
          |||||||||
Db    324 GAAGAGATTATGATGGAGTAATAATTAAGTAAGGATTAATAAGGATTTTGAGTGTGTCAAAA 383

Qy    347 GTATCACACATAATAAAAGTTCTACCTTTCTTAAAAAGTGTAGC 403
          |||||||||
Db    384 AAAACAAAAATTAATAAAAGTGTTGCCTTTCTTTATTACACTTTAATAAGTGGGTAGCTTGTTGAGTTGCAC 443

Qy    404 AATTGGCTTTTTAGTAGCCATTTGACACATTAAAATAAATTTGTGACACATCATT 456
          |||||||
Db    444 TATTGGCCTATAGTAGCCATTTGACACATTAAAATAAATTTGTGACACATCATT 496
```

RESULT 10

```
LOCUS       AI489818                522 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST... tomato ovary, TAMU Lycopersicon esculentum cDNA clone
ACCESSION   AI489818
VERSION     AI489818.1  GI:...
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1..522
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED13B21"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /lab_host="Xl1-Blue MRF'"
                     /clone_lib="tomato ovary, TAMU"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; cLED - Tomato Carpel EST Library. Oligodr-primed and
                     directionally cloned cDNA in vector lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
BASE COUNT      165 a       75 c      104 g      178 t
ORIGIN

Query Match             35.1%;  Score 189.8;  DB 9;  Length 522;
  Best Local Similarity   68.9%;  Pred. No. 2.5e-24;
  Matches  326;  Conservative    0;  Mismatches  127;  Indels   20;  Gaps    4;

Qy      1 ATGGGCTCGCTCCTTCCTTGTGCTTCAATGGCCATTTCTATCCTTGGCCAAGGATGCTCTTCTTTGTTGCC 60
          |||||||||| || || | |||||||| ||||||||
Db     27 ATGGGCTCGTTCCATTTTCTTCATGGCATT                                 TTTGTTACC 86

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTTCTGGATATGC 117
          |||||| ||| | ||||||||   ||||||||
Db     87 TATGAGGTTAGAAGCTTCATCAAATTTGCAAAGCACCAAACTTCCCAGGATTATGT 146

Qy    118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
          |||||||||||
Db    147 TTTATGGACTCATCATGTAGAAAATAITTGTATCAAAGAGAAAATTTACTGGTGGAACATTGT 206

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACACATTTCTGGATATGC 117
Db     87 TATGAGGTTAGAAGCTTCATCAAATTTGCAAAGCACCAAACTTCCCAGGATTATGT 146

Qy    118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGT 177
Db    147 TTTATGGACTCATCATGTAGAAAATAITTGTATCAAAGAGAAAATTTACTGGTGGAACATTGT 206

Qy    178 AGCAAAAATCCTCAGAAGGTGCCTATGTACTTAAGCCACATGTGTGTTTTGATGAGAAGATGACT 237
Db    207 AGCAAACTCCAAAGCAAGAAGTGTCTATGCACTTAAGCCACATGTGTGTTATT---TGACAAAAATCTCA 263

Qy    238 AAAACAGGAGCTGAAATTTGGGCTGAGGAAGCAAAAACTTTGGCCTTGCCTTGAA 297
Db    264 AGTGAAGTTAAAGCAACACTTTGGGTGAGGAAGCAAAAACTCAAGTGAAGTTGTGCTTGAA 323

Qy    298 GAAGAGATAATGGATTAACTAATTAGAGAGATTAGAAGAAAATTAAGGAT 346
Db    324 GAAGAGATTATGATGGAGTAATAATTAAGTAAGGATTAATAAGGATTTTGAGTGTGTCAAAA 383

Qy    347 GTATCACACATAATAAAAGTTCTACCTTTCTTAAAAAGTGTAGC 403
Db    384 AAAACAAAAATTAATAAAAGTGTTGCCTTTCTTTATTACACTTTAATAAGTGGGTAGCTTGTTGAGTTGCAC 443

Qy    404 AATTGGCTTTTTAGTAGCCATTTGACACATTAAAATAAATTTGTGACACATCATT 456
Db    444 TATTGGCCTATAGTAGCCATTTGACACATTAAAATAAATTTGTGACACATCATT 496
```

DEFINITION EST248157 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED15G20, mRNA sequence.
ACCESSION   AI489818
VERSION   AI489818.1 GI:4385189
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 522)
AUTHORS   Alala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J
TITLE   Generation of ESTs from tomato carpel tissue
JOURNAL   Unpublished
COMMENT   Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES   Location/Qualifiers
     source   1..522
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED15G20"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /clone_lib="tomato ovary, TAMU"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     XhoI; cLED - Tomato Carpel EST Library; Oligo dT-primed and
                     directionally cloned cDNA in vector Lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
BASE COUNT   165 a   75 c   104 g   178 t
ORIGIN

Query Match   35.1%; Score 189.8; DB 9; Length 522;
Best Local Similarity   68.9%; Pred. No. 2.5e-24;
Matches 326; Conservative   0; Mismatches 127; Indels 20; Gaps 4;

```
Qy         1 ATGGCTCGCTCCTTGTGCCTTCATGGCCAATTGGTCTATCCTGGCAAGGATGCTCTTTGCC 60
             |||||||||| ||  || ||||||| |||||| ||||||| | |||| |    | ||| ||
Db        27 ATGGCTCGTTCCATTTTCTTCATGGCAATTTTGGTCTTGGGAATAGCTCTTTGTTGTTACC 86

Qy        61 TATGAGGTGCAAGCTAGAGAA---TGCAAAAACAGAAGACAACACATTTCCTGGAATATGC 117
             |||||||| |||||| |  |||   |||||| | | || || ||||| ||| |||| ||||
Db        87 TATGAGGTAGAAGCTCAGCAAAATTGCAAAGGCACCAAGCCAAACTTTCCCAGGATTATGT 146

Qy       118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCGTGAGAAAATTTACTGATGGTCATTGT 177
             ||||| |||||| | ||||||||||||||||||||| | |||||||||||||| || |||||
Db       147 TTTTATGGACTCATCATGTAGAAAATATTGTATCAAAGAGAAATTTACTGTGTTGACATTGT 206

Qy       178 AGCAAAAATCCTCAGAAGGTGCCTATGTGTCGTTGATGAAGAATGACT 237
             ||||| |  |||||||| || | |||||  |  | ||| |  | |||
Db       207 AGCAAACTCCAAAGGAAGTGTCTATGCACTAAGCCATGTGTATT---TGACAAAATCTCA 263

Qy       238 AAAACAGGAGCTGAAATTTGGCTGAGGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAA 297
             ||||||||||||||||||| | | || ||||||||||||| ||| |||||||| ||  | ||
Db       264 AGTGAAGTTAAAGCAACTTTGGGTGAGGAGGAAGCAAAAAACTCTAAGTGAAGTTGTGCTTGAA 323

Qy       298 GAAGAAGAATAATGGATAACTAATTAGAGATTAGAAGAGAAATTAAGGAT------------GCA 346
             |||||  | ||||| |  || ||| || |||| ||||  |||| || |             |||
Db       324 GAAGAGAATTATGAGTGAGTAATAAATTAAGTGTTGCCTTTCTTTATTAGGGTAGCTTGA 383

Qy       347 GTATCACACATAATAAATAAAGTTTCTACCTTTTCTTAAAAGTGTGGTTT 384
             | ||||| ||||||||||  ||| |  | | ||||  | |  || |||  |
Db       384 AAAACAAAATTAATAAAAGTGTTGCCTTTCTTTATTAGGGTAGCTTGATGGTTGTGTTTAG 443

Qy       404 AATTGGCTTTTAGTAGCCTTTTATTACACTTAAATAAGTGTGGCACTTCAAT 456
             ||||||| | ||||||| ||| | || ||| |||||||| ||| |||||| ||
Db       444 TATTGGCCTATTAGTAGCCATTTGACACATTAAATAAGTTGTGACACATCATT 496
```

RESULT 11
LOCUS   AW217379   551 bp   mRNA   linear   EST 18-MAY-2001
DEFINITION EST296063 tomato flower buds 0-3 mm, Cornell University
            Lycopersicon esculentum cDNA clone cTOA1K13 similar to
            gamma-thionin-like protein/acidic protein precursor, mRNA sequence.
ACCESSION   AW217379
VERSION   AW217379.1 GI:6528221
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 551)
AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
            ,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE   Generation of ESTs from tomato flower tissue
JOURNAL   Unpublished
COMMENT   Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES   Location/Qualifiers
     source   1..551
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOA1K13"
                     /tissue_type="flower"
                     /dev_stage="0-3mm buds"
                     /clone_lib="tomato flower buds 0-3 mm, Cornell University"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Tanksley; Flower buds and flowers were
                     taken from greenhouse plants (4-8 wks old, TA496). They
                     were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT   176 a   79 c   107 g   189 t
ORIGIN

Query Match   35.1%; Score 189.8; DB 9; Length 551;
Best Local Similarity   68.9%; Pred. No. 2.4e-24;
Matches 326; Conservative   0; Mismatches 127; Indels 20; Gaps 4;

```
Qy         1 ATGGCTCGCTCCTTGTGTGCTTCATGGCCAATTCTTGGCCAAGGATGCTCTTTGTTGTTGCC 60
             |||||||| ||  ||  | ||||||| ||||||      || |  ||    | ||| |||||
Db        27 ATGGCTCGTTCCATTTTCTTCATGGCAATTTTGGTCTTGGGAATAGCTCTTTGTTGTTACC 86

Qy        61 TATGAGGTGCAAGCTAGAGAA---TGCAAAAACAGAAGACAACACATTTCCTGGAATATGC 117
             |||||||| |||||| |  |||   |||||| | | || || ||||| ||| |||| ||||
Db        87 TATGAGGTAGAAGCTCAGCAAAATTGCAAAGGCACCAAGCCAAACTTTCCCAGGATTATGT 146

Qy       118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCGTGAGAAAATTTACTGATGGTCATTGT 177
             ||||| |||||| | ||||||||||||||||||||| | |||||||||||||| || |||||
Db       147 TTTTATGGACTCATCATGTAGAAAATATTGTATCAAAGAGAAATTTACTGTGTTGACATTGT 206

Qy       178 AGCAAAAATCCTCAGAAGGTGCCTATGTGTCGTTGATGAAGAATGACT 237
             ||||| |  |||||||| || | |||||  |  | ||| |  | |||
Db       207 AGCAAACTCCAAAGGAAGTGTCTATGCACTAAGCCATGTGTATT---TGACAAAATCTCA 263

Qy       238 AAAACAGGAGCTGAAATTTGGCTGAGGAGGAAGCAAAAAACTTTGGCTGCAGCTTGCTTGAA 297
             ||||||||||||||||||| | | || ||||||||||||| ||| |||||||| ||  | ||
Db       264 AGTGAAGTTAAAGCAACTTTGGGTGAGGAGGAAGCAAAAAACTCTAAGTGAAGTTGTGCTTGAA 323
```

RESULT 12

```
LOCUS       AI897456          562 bp    mRNA      linear    EST 18-MAY-2001
DEFINITION  EST266899 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED27F13, mRNA sequence.
ACCESSION   AI897456
VERSION     AI897456.1  GI:5603358
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum
  ORGANISM  Lycopersicon esculentum (tomato)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1..562
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED27F13"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /note="Vector: pBluescript SK(-); Site 1: EcoR1; Site 2:
                     XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
                     directionally cloned cDNA in vector Lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
                     /lab_host="tomato ovary, TAMU"
                     /clone_lib="Xl1-Blue MRF'."
BASE COUNT    180 a    79 c    108 g    194 t    1 others
ORIGIN
```

```
Query Match       35.1%;  Score 189.8;  DB 9;  Length 562;
Best Local Similarity   68.9%;  Pred. No. 2.4e-24;
Matches 326;  Conservative   0;  Mismatches 127;  Indels 20;  Gaps 4;
```

```
Qy       1 ATGGCTCGCTCCTTGTCGTTCATGGCATTGGCAAGGATGCTCTTGTTGTTGTTGCC 60
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db      27 ATGGCTCGCTTCCATTTTCTTCATGGCATTGGCAAGGATGCTCTTGTTCTTGTTACC 86

Qy      61 TATGAGGTGCAAGCTAGAGAA---TGCAAACAGAAAGCAACACATTCCTGGAATTGC 117
           |||||||||||||||||||||    |||||| |||||||||||| |||||| | |||
Db      87 TATGAGGTAGAAGCTCAGCGAAATTGCAAAGCAACAAGCCAAACACTTCCCAGGATTATGT 146

Qy     118 ATTACCAAACCACCATGCAGAGAAAAAGCTTGTATCAGTGAGAAATTACTGGTCATTGT 177
           ||| |||  |||||| ||  ||  | ||||  |   | |||||||| |||| ||||||
Db     147 TTTATGGGACTCATCATGTAGAAAAATATTGTATCAAAAGAAGAAAATTACTGGTGGACATTGT 206
```

RESULT 13

```
LOCUS       AI898009          564 bp    mRNA      linear    EST 18-MAY-2001
DEFINITION  EST267452 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED31M2, mRNA sequence.
ACCESSION   AI898009
VERSION     AI898009.1  GI:5603911
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum
  ORGANISM  Lycopersicon esculentum (tomato)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1..564
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED31M2"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /note="Vector: pBluescript SK(-); Site 1: EcoR1; Site 2:
                     XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
                     directionally cloned cDNA in vector Lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
                     /lab_host="tomato ovary, TAMU"
                     /clone_lib="Xl1-Blue MRF'."
BASE COUNT    185 a    80 c    108 g    190 t    1 others
ORIGIN
```

```
Query Match       35.1%;  Score 189.8;  DB 9;  Length 564;
Best Local Similarity   68.9%;  Pred. No. 2.4e-24;
Matches 326;  Conservative   0;  Mismatches 127;  Indels 20;  Gaps 4;
```

```
Qy       1 ATGGCTCGCTCCTTGTCGTTCATGGCATTGGCAAGGATGCTCTTGTTGTTGTTGCC 60
           |||||| ||| ||  |||| |||||||||||||||||||||| ||||||||||
Db ...
```

(Sequence alignment rows)

```
Qy     298 GAAGAGGATAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGAT---------------GCA 346
           |||||||||||||||||||||||||||||||||||||||||||||||               |||
Db     324 GAAGAGGATTATGATGGAGTAATAATTAAGTTAGGAGGTTAAATAGGATTTGAGTGGTGTTGTTCAAAA 383

Qy     347 GTATCACACAATAATAAAAGTTTCTACCTTTCTTAAAAAGTGTAGC--TAATGTTGTGTTTT 403
           ||||||||||||||||||||||| ||||||||||||||||||||    |||||||||||||
Db     384 AAAACAGAAAATTAATAAAAGTGTGCCCTTTTCTTAATTAGGGTAGCTTGTGATGTTGTGTTAG 443

Qy     404 AATTGGCTTTTAGTTAGTTAGCCTTTTATTACACTTTAATAAGTGTGTGACACATTAATAAATAAAGTTGTGACACATCATT 456
           ||||||||||||||||||||||||||||||
Db     444 TATTGGCCTATAGTAGCCATTTGACACATTAATAAATAAAGTTGTGACACATCATT 496
```

RESULT 14

```
Db    27 ATGGCTCGTTCCATTTCTTCCATGGCCATTTGCTTGGTCTTGGCCAATGATGATGCTTTGTTACC
Qy       |||||||||||||||||||||||||||||||||||||||||
Qy    61 TATGAGGTTGCAAGCTAGAGAA---TTGCAAAAACAGAAAGCAACACACATTTCCTTGGAATATGC 117
Db       ||||||||||||||||||||||||||||||||||||||||
Db    87 TATGAGGTAGAAGCTCAGCAAAATTTGCAAAGCCAACAAAACTTTCCCAGGATTATGT 146
Qy       |||||||||||||||||||||||||||||||
Qy   118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTCATCGTGAGAAATTTACTGATGGTCATTGT 177
Db       |||||||||||||||||||||||||||||||||
Db   147 TTTATGGACTCATCGTAGAAAAATATTGTATCAAAGACAGAAAATTTACTGGTGGACATTGT 206
Qy       |||||||||||||||||||||||||||||||
Qy   178 AGCAAAAATCCTCAGAAGGTGCCTATCGTGAGAAGCCATGTGTGTTTGATGAGAAGATGACT 237
Db       ||||||||||||||||||||||||||||||
Db   207 AGCAAAACTCCAAAGGAAGTGTCTATCGTCCACTAAGCCATGTGTATT---TGACAAAATCTCA 263
Qy       |||||||||||||||||||
Qy   238 AAAACAGGAGCTGAAATTTGGCTGAAGAAGCAAAAACTTGGCCTCAGCTTTGCTTGAA 297
Db       |||||||||||||||||||||||||||||||
Db   264 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTTGCTTGAA 323
Qy       ||||||||||||||||||||||||||||
Qy   298 GAAGAGATAATGGATAACTAATTAGGAGATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 346
Db       ||||||||||||||||||||||||   ------GCA
Db   324 GAAGAGATTATGATGGAGTAATAATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 383
Qy       ||||||||||||||||||||||||||||||
Qy   347 GTATCACACATAATAAATAAAGTTTCTTACCTTTCTTAAAAGTGTGTAGC--TAATGTTGTGTTTT 403
Db       ||||||||||||||||||||||||||||
Db   384 AAAACAAAATTAATAAAAGTGTTGCCTTTTCTTTATTAGGGAGGTTAGCTTGTGAIGTTGTGTTTAG 443
Qy       ||||||||||||||||||||||
Qy   404 AAATTGGCTTTTAGTAGCCTTTTATTTACACTTTAAATAAGTTAAATAAGTTTGTGACACATCAAT 456
Db       |||||||||||||||||||||||||||||||
Db   444 TATTGGCCTATAGTAGCCATTTGACACATTAAATAAGTTTGTGACACATCATT 496
```

LOCUS   BI931665
DEFINITION EST551554 tomato flower, 8 mm to preanthesis buds Lycopersicon
           esculentum cDNA clone cTOC20P19 5' end, mRNA sequence.
ACCESSION BI931665
VERSION   BI931665.1 GI:16246137
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 712)
AUTHORS   Van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
           Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
           Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE   Generation of ESTs from tomato flower tissue, buds 8 mm -
           preanthesis
JOURNAL   Unpublished
COMMENT   Contact: CUGI
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Email: http://www.genome.clemson.edu/orders/index.html
           This clone is available through the Clemson University Genomics
           Institute
           Seq primer: T3.
FEATURES   Location/Qualifiers
     source   1..712
                    /organism="Lycopersicon esculentum"
                    /mol_type="mRNA"
                    /cultivar="TA496"
                    /db_xref="taxon:4081"
                    /clone="cTOC20P19"
                    /tissue_type="flower"
                    /clone_lib="buds 8mm to preanthesis"
                    /dev_stage="buds 8 mm to preanthesis buds"
                    /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                    XhoI; supplier: Cornell University; sequencing: The
                    Institute for Genomic Research; Flower buds and flowers
                    were taken from greenhouse plants (4-8 wks old, TA496)."

LOCUS   BI931665   712 bp   mRNA   linear   EST 18-OCT-2001

BASE COUNT   235 a   114 c   125 g   238 t
ORIGIN

"They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

Query-Match   35.1%;   Score 189.8;   DB 12;   Length 712;
Best Local Similarity 68.9%;   Pred. No. 2.2e-24;
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

RESULT 15

```
Qy     1 ATGGCTCGTTCCTTCCTTTGCTTCATGGCCATTTGCTTGGTCTTCTTTGGTTGCC 60
Db       |||||||||||||||||||||||||||||||||
Db    58 ATGGGTCGTTCCATTTTCTTCCATGGCCATTTGCTTGGTCTTCTTTGGTTCC
Qy    61 TATGAGGTTGCAAGCTAGAGAA---TTGCAAAAACAGAAAGCAACACACATTTCCTTGGAATATGC 117
Db       ||||||||||||||||||||||||||||||||||
Db    58 ATGGGTCGTTCCATTTTCTTCCATGGCCATTTGCTTGGTCTTCTTTGGTTCC 117
Qy   118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTCATCGTGAGAAATTTACTGATGGTCATTGT 177
Db       |||||||||||||||||||||||||||||||
Db   178 TTTATGGACTCATCGTAGAAAAATATTGTATCAAAGACAGAAAATTTACTGGTGGACATTGT 237
Qy   178 AGCAAAAATCCTCAGAAGGTGCCTATCGTGAGAAGCCATGTGTGTTTGATGAGAAGATGACT 237
Db       |||||||||||||||||||||||||||||||
Db   238 AGCAAAACTCCAAAGGAAGTGTCTATCGTCCACTAAGCCATGTGTATT---TGACAAAATCTCA 294
Qy   238 AAAACAGGAGCTGAAATTTGGCTGAAGAAGCAAAAACTTGGCCTCAGCTTTGCTTGAA 297
Db       |||||||||||||||||||||||||||||||
Db   295 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTTGCTTGAA 354
Qy   298 GAAGAGATAATGGATAACTAATTAGGAGATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 346
Db       ||||||||||||||||||||||||   ------GCA
Db   355 GAAGAGATTATGATGGAGTAATAATTAAGTGAGGTTAAATAAGGATTTTGAGTGTCAAAA 414
Qy   347 GTATCACACATAATAAATAAAGTTTCTTACCTTTCTTAAAAGTGTGTAGC--TAATGTTGTGTTTT 403
Db       |||||||||||||||||||||||||||||||
Db   415 AAAACAAAATTAATAAAAGTGTTGCCTTTTCTTTATTAGGGAGGTTAGCTTGTGAIGTTGTGTTTAG 474
Qy   404 AAATTGGCTTTTAGTAGCCTTTTATTACACTTTAAATAAGTTAAATAAGTTTGTGACACATCAAT 456
Db       ||||||||||||||||||||||||||||||
Db   475 TATTGGCCTATAGTAGCCATTTGACACATTAAATAAGTTTGTGACACATCATT 527
```

LOCUS   AW623541
DEFINITION EST321486 tomato flower buds 3-8 mm, Cornell University
           Lycopersicon esculentum cDNA clone cTOB11C9 5', mRNA sequence.
ACCESSION AW623541
VERSION   AW623541.1 GI:7336568
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 500)
AUTHORS   van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
           ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
           Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE   Generation of ESTs from tomato flower tissue, 3-8 mm buds
JOURNAL   Unpublished
COMMENT   Contact: CUGI
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Email: http://www.genome.clemson.edu/orders/index.html
           5 prime sequence.
FEATURES   Location/Qualifiers
     source   1..500
                    /organism="Lycopersicon esculentum"
                    /mol_type="mRNA"
                    /cultivar="TA496"
                    /db_xref="taxon:4081"

LOCUS   AW623541   500 bp   mRNA   linear   EST 18-MAY-2001

```
                         /clone="cTOB11C9"
                         /tissue_type="flower"
                         /dev_stage="3-8mm buds"
                         /clone_lib="tomato flower buds 3-8 mm, Cornell University"
                         /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                         XhoI; supplier: Tanksley; Flower buds and flowers were
                         taken from greenhouse plants (4-8 wks old, TA496). They
                         were immediately frozen in liquid nitrogen and then
                         size-separated while remaining frozen."

BASE COUNT       160 a        77 c       105 g       158 t
ORIGIN

Query Match      34.8%;       Score 188.2;  DB 9;   Length 500;
Best Local Similarity  68.7%;  Pred. No. 4.8e-24;
Matches 307;  Conservative    0;  Mismatches 128;  Indels  12;  Gaps   3;


Qy      1 ATGGCTCGCTCCCTTGTGCTTCATGGCCATTTGCTATCTTGGCAAGGATGGCTTCTTTGCC 60
          ||||||||| || ||| | ||||||||||||| | ||||||| |||||| ||| |||||||
Db     34 ATGGCTCCGTTCATTTGCTTCATGGCACTTATGGTCTTGGGCAATGGTGCTCTTTGTTTCC 93
        .

Qy     61 TATGAGGTGCAAGCTAGAGA---ATGCAAAACAGAAAAGCAACACATTTCCTGGAATATGC 117
          |||||||||||| || |||||    |||||||| |||||| ||| | ||| ||||||||||
Db     94 TCTGAGGTGCAAGCTCAACAGATGTGCAAATCAACAAGCCAAAACCTTCAAGGGAATTATGC 153

Qy    118 ATTACCAAAACCACCATGCAGAAAAGCTTGTTATCAGTGAGAAAATTTACTGATGGTCATTGT 177
          |||| |||| ||| |||||| |||| |||||||| |||||||||||| |||||||||||||||
Db    154 TTTACCGATTCATCGTGTGATAAAAGGCTTGTTGTCACACAGGAGTTTACACGGTGGRACATTGT 213

Qy    178 AGCAAAAATCCTCAGAAGGTGGCCTATGTACTAAGCCAATGTGTGTTTGATGAGAAGAATGACT 237
          | ||||| |||||||||||||||| |||| ||| |||||||||| |||||||||||||||||||
Db    214 AGCAAAACTCCAAAGAAGAGTGCCTATGTACTAAGGTTTTGTATTTGAAAAAGAATTCAAAT 273

Qy    238 AAAACAGGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTTGAA 297
          ||   | |  | |||||||||| |||||||| ||||| ||| |||||| |||||||||||||
Db    274 GAAGTTAAAAC---AACTTTGGTTGGGGAAGCAAAAACTCTAAGTGAAACTGTGCTTTGAA 330

Qy    298 GAAGAGATAATGGATAACTAATTA------GAGATTAGAGAAAATTAAGGATGGCAGTATC 351
          ||||||||| ||||||||||||||      |||||||| || ||||||||||||| || ||
Db    331 GAAGAGAATTGTGATGGAGTAATTAATTAAGTGAGATAATTAAGAATTTGAGTGTCAAAAATC 390

Qy    352 ACACATAATAAAGTTTCTACCCTTTCTTCATTAAAGTGTAGGTAAGTTGTGTTTTAATTGGCT 411
          || | |||| ||||| | |||  | | ||| |||||||| ||| || | ||| | |||||||||
Db    391 AAAATTAATAAAGTATTGTCTTTTAGTAAAAAGGGTAGCTTCTAAGTGTCGGTGTTTAATTGGGC 450

Qy    412 TTTAGTAGCCTTTTTATTTACACTTTAAA 438
          || ||||||||| |  |||| | ||||||
Db    451 CTTAGTAGCCATTTGATACAATTAAATA 477


Search completed: January 28, 2004, 09:54:53
Job time : 1952 secs
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 08:46:51

Title: US-10-072-809B-18
Perfect score: 552
Sequence: 1 MARSLCFMAFAILARMLFVA.........LAEEAKTLAAALEEEIMDN 105

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

; Search time 21 Seconds
(without alignments)
211.554 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	333	60.3	188	6	5175095-3	Patent No.5175095
2	188	34.1	83	4	US-09-442-631-2	Sequence 2, Appli
3	122.5	22.2	74	4	US-09-442-631-4	Sequence 4, Appli
4	121	21.9	47	4	US-08-377-687-28	Sequence 28, Appl
5	121	21.9	47	1	US-08-656-318A-13	Sequence 13, Appl
6	121	21.9	47	2	US-08-777-192-28	Sequence 28, Appl
7	121	21.9	47	3	US-08-955-459-13	Sequence 13, Appl
8	121	21.9	47	3	US-08-971-988-28	Sequence 28, Appl
9	121	21.9	47	2	US-08-289-458-2	Sequence 2, Appli
10	120	21.7	75	2	US-08-761-549-2	Sequence 2, Appli
11	120	21.7	75	3	US-08-127-646-2	Sequence 2, Appli
12	120	21.7	75	2	US-08-377-687-49	Sequence 49, Appl
13	110	19.9	80	2	US-08-777-192-49	Sequence 49, Appl
14	110	19.9	80	3	US-08-971-988-49	Sequence 49, Appl
15	110	19.9	80	1	US-09-103-489-20	Sequence 20, Appl
16	108.5	19.7	79	1	US-09-103-489-15	Sequence 15, Appl
17	108.5	19.7	79	3	US-08-627-706-15	Sequence 15, Appl
18	108	19.6	80	2	US-08-377-687-59	Sequence 59, Appl
19	108	19.6	80	2	US-08-777-192-59	Sequence 59, Appl
20	108	19.6	80	3	US-08-971-988-59	Sequence 59, Appl
21	107	19.4	74	4	US-08-543-238-5	Sequence 5, Appli
22	107	19.4	74	1	US-08-420-526-5	Sequence 5, Appli
23	99	17.9	48	1	US-08-377-687-29	Sequence 29, Appl
24	99	17.9	48	1	US-08-656-318A-11	Sequence 11, Appl
25	99	17.9	48	2	US-08-777-192-29	Sequence 29, Appl
26	99	17.9	48	2	US-08-955-459-11	Sequence 11, Appl
27	99	17.9	48	3	US-08-971-982-29	Sequence 29, Appl
28	97	17.6	47	4	US-09-077-951-21	Sequence 21, Appl
29	95	17.2	74	1	US-08-543-238-2	Sequence 2, Appli
30	95	17.2	74	1	US-08-420-526-2	Sequence 2, Appli
31	90.5	16.4	51	1	US-09-077-951-61	Sequence 61, Appl
32	88.5	16.0	51	4	US-09-077-951-25	Sequence 25, Appl
33	88.5	16.0	51	2	US-09-077-951-46	Sequence 46, Appl
34	88.5	16.0	51	4	US-09-077-951-66	Sequence 66, Appl
35	88.5	16.0	51	4	US-09-077-951-76	Sequence 76, Appl
36	88	16.0	47	1	US-08-656-318A-12	Sequence 12, Appl
37	88	15.9	47	2	US-08-377-687-30	Sequence 30, Appl
38	88	15.9	47	2	US-08-777-192-30	Sequence 30, Appl
39	88	15.9	47	3	US-08-955-459-12	Sequence 12, Appl
40	88	15.9	51	2	US-09-077-951-26	Sequence 26, Appl
41	86.5	15.9	51	4	US-08-971-982-30	Sequence 30, Appl
42	86.5	15.7	72	4	US-09-003-198A-17	Sequence 17, Appl
43	85.5	15.5	46	3	US-08-632-511A-5	Sequence 5, Appli
44	85.5	15.5	46	3	US-08-091-590A-11	Sequence 11, Appl
45	85.5	15.5	46	4	US-09-488-200-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
5175095-3
;Patent No.5175095
; APPLICANT: Martineau, Belinda M.;Houck, Catherine M.
; TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/554,195
; FILING DATE: 17-JUL-1990
;SEQ ID NO:3:
; LENGTH: 188
5175095-3

Query Match   60.3%; Score 333; DB 6; Length 188;
Best Local Similarity   62.5%; Pred. No. 1e-30;
Matches   65; Conservative   12; Mismatches   25; Indels   2; Gaps   2;

```
Qy    1 MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
          ||||  |||| ||  |||||||     ||  ||  |||| |     ||| |  ||| |||
Db   14 MARSIFFMAFLVLAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC 73

Qy   60 SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALEEEIM 103
          || ||||||||||||| |  ||  | ||||| | |   |||||
Db   74 SKLQRKCLCTKPCVFD-KISSEVRATLGERAKTLSEVVLEEEIM 116
```

RESULT 2
US-09-442-631-2
; Sequence 2, Application US/09442631
; Patent No. 6300489
; GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-442-631-2

Query Match   34.1%; Score 188; DB 4; Length 83;

Best Local Similarity 45.5%; Conservative 45; Pred No. 1.7e-14;
Matches 45; Conservative 9; Mismatches 27; Indels 18; Gaps 3;

```
Qy    1 MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACI-SEKFTDGHC 59
        ||||:|| ||:|| :||      |:||:    |::|::  |:  :::   |||:  ||||
Db    1 MARSIYFMAFLVLA-TLFVAYGVQGKEICCKELTKPVKCSSDPLCQKLCMEKEKYEDGHC 59

Qy   60 SKILRRCLCTKPCVFDEKMTKTGAEILAEBAKTLAAAII 98
        :|| | |:|||
Db   60 FTILSKCLCMKRC----------NAKTLATELL 82
```

RESULT 3
US-09-442-631-4
Sequence 4, Application US/09442631
Patent No. 6300489
GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-442-631-4

Query Match 22.2%; Score 122.5; DB 4; Length 74;
Best Local Similarity 32.9%; Pred. No. 4.9e-07;
Matches 23; Conservative 12; Mismatches 28; Indels 7; Gaps 1;

```
Qy   10 FAILARMLFVAYEV-----QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKI 62
        :||:  |||: ||      |:   |: ||||   || |   ||||||||||
Db    4 FSKVVATFLMKVFATDMMAEAKICEALSGNFKGLCLSSRDCGNVCRREGFTDGSCIGF 63
```

```
Qy   63 LRRCLCTKPC 72
        |||:|||
Db   64 RLQCFCTKPC 73
```

RESULT 4
US-08-377-687-28
Sequence 28, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 800
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-377-687-28

Query Match 21.9%; Score 121; DB 1; Length 47;
Best Local Similarity 42.6%; Pred. No. 4.4e-07;
Matches 20; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

```
Qy   26 RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
        |:||:|||   ||:|:|||   |:|  | || |||  ||:|||||
Db    1 RHCESLSHRFKGPCTRDSNCASVCETERFSGGNCHGFRRRCFCTKPC 47
```

RESULT 5
US-08-656-318A-13
Sequence 13, Application US/08656318A
; Patent No. 5750504
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,318A
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326424.0
; FILING DATE: 24-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/02766
; FILING DATE: 19-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 224199/SEE37925/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3075
```

TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 47 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   ORIGINAL SOURCE:
;     ORGANISM: p322
US-08-656-318A-13

Qy   26  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCCTKPC  72
         |||:|:|  |:| .: |:.  :  | |  | |.   |||:|||  ||||
Db    1  RHCESLSHRFKGPCTRDSNCASVCETERFSGGNCHGFRRRCFCTKPC  47

Query Match          21.9%;   Score 121;  DB 1;  Length 47;
Best Local Similarity  42.6%;   Pred. No. 4.4e-07;
Matches  20;  Conservative  7;  Mismatches  20;  Indels  0;  Gaps  0;

RESULT 6
US-08-777-192-28
Sequence 28, Application US/08777192
Patent No. 5824869
; GENERAL INFORMATION:
;   APPLICANT: BROEKAERT, WILLEM F.
;   APPLICANT: CAMMUE, BRUNO P.A.
;   APPLICANT: OSBORN, RUPERT W.
;   APPLICANT: REES, SARAH B.
;   APPLICANT: TERRAS, FRANKY R.G.
;   APPLICANT: VANDERLEYDEN, JOZEF
;   TITLE OF INVENTION: BIOCIDAL PROTEINS
;   NUMBER OF SEQUENCES: 59
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: CUSHMAN DARBY & CUSHMAN
;     STREET: 1100 NEW YORK AVENUE, N.W.
;     CITY: WASHINGTON
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/777,192
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/002,480
;     FILING DATE: 04-JAN-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: KOKULIS, PAUL N.
;     REGISTRATION NUMBER: 16,773
;     REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-861-3000
;     TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 28:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 47 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-777-192-28

Query Match          21.9%;   Score 121;  DB 2;  Length 47;
Best Local Similarity  42.6%;   Pred. No. 4.4e-07;
Matches  20;  Conservative  7;  Mismatches  20;  Indels  0;  Gaps  0;

Qy   26  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCCTKPC  72
         |||:|:|  |:| .: |:.  :  | |  | |.   |||:|||  ||||
Db    1  RHCESLSHRFKGPCTRDSNCASVCETERFSGGNCHGFRRRCFCTKPC  47

RESULT 7
US-08-956-459-13
Sequence 13, Application US/08956459
Patent No. 5919918
; GENERAL INFORMATION:
;   APPLICANT: BROEKAERT, WILLEM F.
;   APPLICANT: CAMMUE, BRUNO P.A.
;   APPLICANT: OSBORN, RUPERT W.
;   APPLICANT: REES, SARAH B.
;   TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
;   NUMBER OF SEQUENCES: 13
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
;     STREET: 1100 New York Avenue, N.W.
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: U.S.A.
;     ZIP: 20005-3918
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Microsoft Word
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/956,459
;     FILING DATE: 22-OCT-1996
;     CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/656,318
;     FILING DATE: 12-JUN-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/GB94/02766
;     FILING DATE: 19-DEC-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: GB 9326424.0.
;     FILING DATE: 24-DEC-1993
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 47 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   ORIGINAL SOURCE:
;     ORGANISM: p322
US-08-956-459-13

Qy   26  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCCTKPC  72
         |||:|:|  |:| .: |:.  :  | |  | |.   |||:|||  ||||
Db    1  RHCESLSHRFKGPCTRDSNCASVCETERFSGGNCHGFRRRCFCTKPC  47

Query Match          21.9%;   Score 121;  DB 2;  Length 47;
Best Local Similarity  42.6%;   Pred. No. 4.4e-07;
Matches  20;  Conservative  7;  Mismatches  20;  Indels  0;  Gaps  0;
US-08-971-982-28

RESULT 8
US-08-971-982-28
Sequence 28, Application US/08971982
Patent No. 6187904
; GENERAL INFORMATION:
;   APPLICANT: BROEKAERT, WILLEM F.
;   CAMMUE, BRUNO P.A.
;   OSBORN, RUPERT W.
;   REES, SARAH B.
;   TERRAS, FRANKY R.G.
;   VANDERLEYDEN, JOZEF
;   TITLE OF INVENTION: BIOCIDAL PROTEINS

NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,982
FILING DATE: 17-No. 6187904-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/002,480
FILING DATE: 04-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-08-971-982-28

Query Match   21.9%;  Score 121;  DB 3;  Length 47;
Best Local Similarity   42.6%;  Pred. No. 4.4e-07;
Matches  20;  Conservative  7;  Mismatches  20;  Indels  0;  Gaps  0;

```
Qy   26 RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRCLCTKPC 72
Db    1 RHCESLSHRFKGPCTRDSNCASVCETERFSGNCHGFRRRCFCTKPC 47
```

RESULT 9
US-08-289-458-2
; Sequence 2, Application US/08289458
; Patent No. 5608144
; GENERAL INFORMATION:
; APPLICANT: BADEN, Catherine S., DUNSMUIR, Pamela,
; APPLICANT: LEE, Kathleen Y.
; TITLE OF INVENTION: PLANT Gp2 PROMOTERS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,458
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12176-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-289-458-2

Query Match   21.7%; Score 120; DB 1; Length 75;
Best Local Similarity   34.3%; Pred. No. 9.7e-07;
Matches 23; Conservative 12; Mismatches 30; Indels 2; Gaps 1;

```
Qy    8 MAFAILARMLFVAYEV--QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRR 65
Db    8 VAIIFLMMLLVFAIDWAAEAKICEALTGNFKGLCLSSRDCGNVCRREGFTDGSCIGFRLQ 67

Qy   66 CLCTKPC 72
Db   68 CFCTKPC 74
```

RESULT 10
US-08-761-549-2
; Sequence 2, Application US/08761549
; Patent No. 5981727
; GENERAL INFORMATION:
; APPLICANT: BADEN, Catherine S., DUNSMUIR, Pamela,
; APPLICANT: LEE, Kathleen Y.
; TITLE OF INVENTION: PLANT Gp2 PROMOTERS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/289,458
; FILING DATE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,549
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12176-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-761-549-2

Query Match   21.7%; Score 120; DB 2; Length 75;
Best Local Similarity   34.3%; Pred. No. 9.7e-07;
Matches 23; Conservative 12; Mismatches 30; Indels 2; Gaps 1;

Qy   8 MAFAILARMLFVAYEV--QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRR 65
              :|: ||..:. : |: .:||||.||||||||:||||||
Db   8 VATIFLWMLLVFATDMWAEAKICEALTGNFKGLCLSSRDCGNVCRREGFTDSSCIGPRLQ 67

Qy   66 CLCTKPC 72
           | ||||||
Db   68 CFCTKPC 74

RESULT 11
US-09-127-646-2
Sequence 2, Application US/09127646
Patent No. 6291744
GENERAL INFORMATION:
  APPLICANT: Baden, Catherine S.
  APPLICANT: Dunsmuir, Pamela
  APPLICANT: Lee, Kathleen Y.
  APPLICANT: DNA Plant Technology Corporation
  TITLE OF INVENTION: Nucleic Acids Encoding Plant Group 2 Proteins and Uses
  TITLE OF INVENTION: Thereof
  FILE REFERENCE: 012176-00402US
  CURRENT APPLICATION NUMBER: US/09/127,646
  CURRENT FILING DATE: 1998-07-31
  EARLIER APPLICATION NUMBER: US 08/289,458
  EARLIER FILING DATE: 1994-08-12
  EARLIER APPLICATION NUMBER: US 08/761,549
  EARLIER FILING DATE: 1996-12-06
  NUMBER OF SEQ ID NOS: 12
  SOFTWARE: PatentIn Ver. 2.0
  SEQ ID NO 2
  LENGTH: 75
  TYPE: PRT
  ORGANISM: Capsicum annuum

Query Match   21.7%; Score 120; DB 3; Length 75;
Best Local Similarity   34.3%; Pred. No. 9.7e-07;
Matches   23; Conservative   12; Mismatches   30; Indels   2; Gaps   1;

Qy   8 MAFAILARMLFVAYEV--QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRR 65
              :|: ||..:. : |: .:||||.||||||||:||||||
Db   8 VATIFLWMLLVFATDMWAEAKICEALTGNFKGLCLSSRDCGNVCRREGFTDSSCIGPRLQ 67

Qy   66 CLCTKPC 72
           | ||||||
Db   68 CFCTKPC 74

RESULT 12
US-08-377-687-49
Sequence 49, Application US/08377687
Patent No. 5538525
GENERAL INFORMATION:
  APPLICANT: BROEKAERT, WILLEM F.
  APPLICANT: CAMMUE, BRUNO P.A.
  APPLICANT: OSBORN, RUPERT W.
  APPLICANT: REES, SARAH B.
  APPLICANT: TERRAS, FRANKY R.G.
  APPLICANT: VANDERLEYDEN, JOZEF
  TITLE OF INVENTION: BIOCIDAL PROTEINS
  NUMBER OF SEQUENCES: 59
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: CUSHMAN DARBY & CUSHMAN
  STREET: 1100 NEW YORK AVENUE, N.W.
  CITY: WASHINGTON
  STATE: D.C.
  COUNTRY: USA
  ZIP: 20005
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/377,687
  FILING DATE:
  CLASSIFICATION: 800
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/002,480
  FILING DATE: 04-JAN-1993
  ATTORNEY/AGENT INFORMATION:
  NAME: KOKULIS, PAUL N.
  REGISTRATION NUMBER: 16,773
  REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 202-861-3000
  TELEFAX: 202-822-0944
  INFORMATION FOR SEQ ID NO: 49:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 80 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-377-687-49

Query Match   19.9%; Score 110; DB 1; Length 80;
Best Local Similarity   32.5%; Pred. No. 1.5e-05;
Matches   26; Conservative   16; Mismatches   30; Indels   8; Gaps   4;

Qy   1 MARSLCFMAFAILARMLFVAYE----VQARE-CKTESNTFPGICITKPPCRKACIS-EKF 54
            |:: ... ||.|||||. :.|| : :.: .|: :||||.||||||||:||||||: ||:
Db   1 MAKFASIIALLFAALVLFAAFEAPTMVEAQKLCERPSGTWSGVCGNNNACKNQCINLEKA 60

Qy   55 TDGHCSKIL--RRCLCTKPC 72
            :||||:|| |:| ||||||
Db   61 RHGSCNYVFPAHKCICYFPC 80

RESULT 13
US-08-777-192-49
Sequence 49, Application US/08777192
Patent No. 5824869
GENERAL INFORMATION:
  APPLICANT: BROEKAERT, WILLEM F.
  APPLICANT: CAMMUE, BRUNO P.A.
  APPLICANT: OSBORN, RUPERT W.
  APPLICANT: REES, SARAH B.
  APPLICANT: TERRAS, FRANKY R.G.
  APPLICANT: VANDERLEYDEN, JOZEF
  TITLE OF INVENTION: BIOCIDAL PROTEINS
  NUMBER OF SEQUENCES: 59
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: CUSHMAN DARBY & CUSHMAN
  STREET: 1100 NEW YORK AVENUE, N.W.
  CITY: WASHINGTON
  STATE: D.C.
  COUNTRY: USA
  ZIP: 20005
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/777,192
  FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/002,480
  FILING DATE: 04-JAN-1993
  ATTORNEY/AGENT INFORMATION:
  NAME: KOKULIS, PAUL N.
  REGISTRATION NUMBER: 16,773
  REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 202-861-3000

TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-777-192-49

Query Match   19.9%;   Score 110;   DB 2;   Length 80;
Best Local Similarity   32.5%;   Pred. No. 1.5e-05;
Matches 26;   Conservative 16;   Mismatches 30;   Indels 8;   Gaps 4;

```
Qy  1 MARSLCFMAFAILARMLFVAYE-----VQARE-CKTESNTFPGICITKPPCRKACIS-EKF 54
        ||:  |:   ||  |::|  |:|   :|:  ||| | ||:|
Db  1 MAKFASIIALLFAALVLFAAFEAPTMVEAQKLCERPSGTWSGVCGNNNACKNQCINLEKA 60

Qy 55 TDGHCSKIL--RRCLCTKPC 72
        |:|  ||:  ||  |||
Db 61 RHGSCNYVFPAHKCICYFPC 80
```

RESULT 14
US-08-971-982-49
Sequence 49, Application US/0897,982
Patent No. 6187904
GENERAL INFORMATION:
APPLICANT: BROEKAERT, WILLEM F.
CAMMUE, BRUNO P.A.
OSBORN, RUPERT W.
REES, SARAH B.
TERRAS, FRANKY R.G.
VANDERLEYDEN, JOZEF
TITLE OF INVENTION: BIOCIDAL PROTEINS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 2005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,982
FILING DATE: 17-No. 6187904-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/002,480
FILING DATE: 04-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-971-982-49

Query Match   19.9%;   Score 110;   DB 3;   Length 80;
Best Local Similarity   32.5%;   Pred. No. 1.5e-05;
Matches 26;   Conservative 16;   Mismatches 30;   Indels 8;   Gaps 4;

```
Qy  1 MARSLCFMAFAILARMLFVAYE-----VQARE-CKTESNTFPGICITKPPCRKACIS-EKF 54
        ||:  |:   ||  |::|  |:|   :|:  ||| | ||:|
Db  1 MAKFASIIALLFAALVLFAAFEAPTMVEAQKLCERPSGTWSGVCGNNNACKNQCINLEKA 60

Qy 55 TDGHCSKIL--RRCLCTKPC 72
        |:|  ||:  ||  |||
Db 61 RHGSCNYVFPAHKCICYFPC 80
```

RESULT 15
US-09-103-489-20
Sequence 20, Application US/09103489
Patent No. 6215048
GENERAL INFORMATION:
APPLICANT: Liang, Jihong
APPLICANT: Shah, Dilip M.
APPLICANT: Wu, Yonnie S.
APPLICANT: Rosenberger, Cindy A.
TITLE OF INVENTION: Antifungal Polypeptide and Methods for
TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
STREET: 700 Chesterfield Village Parkway No. 6215048th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,489
FILING DATE: 24-JUN-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Cohen, Charles E.
REGISTRATION NUMBER: 34,565
REFERENCE/DOCKET NUMBER: 38-21 (10700)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-6224
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-103-489-20

Query Match   19.9%;   Score 110;   DB 3;   Length 80;
Best Local Similarity   32.5%;   Pred. No. 1.5e-05;
Matches 26;   Conservative 16;   Mismatches 30;   Indels 8;   Gaps 4;

```
Qy  1 MARSLCFMAFAILARMLFVAYE-----VQARE-CKTESNTFPGICITKPPCRKACIS-EKF 54
        ||:  |:   ||  |::|  |:|   :|:  ||| | ||:|
Db  1 MAKFASIIALLFAALVLFAAFEAPTMVEAQKLCERPSGTWSGVCGNNNACKNQCINLEKA 60

Qy 55 TDGHCSKIL--RRCLCTKPC 72
        |:|  ||:  ||  |||
Db 61 RHGSCNYVFPAHKCICYFPC 80
```

Search completed: January 28, 2004, 08:49:14
Job time : 22 secs

```
OM protein - protein search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:             January 28, 2004, 08:48:46 ; Search time 371 Seconds
                                                 (without alignments)
                                                 58.510 Million cell updates/sec

Title:              US-10-072-809B-18
Perfect score:      552
Sequence:           1 MARSLCFWAFAILARMLFVA........LAEEAKTLAAALLEEEIMDN 105

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters:   777136

Post-processing: Minimum Match 0%
                 Maximum Match 100%
Minimum DB seq length: 0             Listing first 45 summaries
Maximum DB seq length: 200000000

Database :  Published Applications AA:*
  1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
  2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
  3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
  4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
  5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
  6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
  7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
  8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
  9: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```
Result              %
No.   Score   Query Match   Length   DB   ID              Description
----  -----   -----------   ------   --   --              -----------
  1    552     100.0         105     12   US-10-072-809A-18   Sequence 18, Appli
  2    529      95.8         105     12   US-10-072-809A-20   Sequence 20, Appl
  3    508      92.0         105     12   US-10-072-809A-51   Sequence 51, Appl
  4    468.5    84.9         106     12   US-10-072-809A-52   Sequence 52, Appl
  5    431      78.1         106     12   US-10-072-809A-16   Sequence 16, Appl
  6    395      71.6          80     12   US-10-072-809A-14   Sequence 14, Appl
  7    342      62.0          72     12   US-10-072-809A-50   Sequence 50, Appl
  8    333      60.3          79     12   US-10-072-809A-21   Sequence 21, Appl
  9    274      49.6         105     12   US-10-072-809A-8    Sequence 8, Appli
 10    272      49.3          47     12   US-10-072-809A-26   Sequence 26, Appl
 11    185      33.5          47     12   US-10-072-809A-25   Sequence 25, Appl
 12    163      29.5          78     15   US-10-178-213-278   Sequence 278, App
 13    160      29.0          75     15   US-10-178-213-290   Sequence 290, App
 14    157      28.4          33     12   US-10-072-809A-12   Sequence 12, Appl
 15    155      28.1          78     15   US-10-178-213-431   Sequence 431, App
 16    151      27.4          74     15   US-10-178-213-2     Sequence 2, Appli
 17    150.5    27.3          77     15   US-10-178-213-407   Sequence 407, App
 18    148      26.8          72     15   US-10-178-213-413   Sequence 413, App
 19    148      26.8          72     15   US-10-178-213-83    Sequence 83, Appli
 20    147      26.6          84     15   US-10-178-213-386   Sequence 386, App
 21    145.5    26.4          79     15   US-10-178-213-293   Sequence 293, App
 22    145      26.3          47     15   US-10-178-213-279   Sequence 279, App
 23    145      26.3          78     15   US-10-072-809A-23   Sequence 86, Appl
 24    144      26.1          78     12   US-10-072-809A-22   Sequence 22, Appl
 25    143      25.9          78     15   US-10-178-213-443   Sequence 443, App
 26    141      25.5          78     15   US-10-178-213-308   Sequence 308, App
 27    140.5    25.5          77     15   US-10-178-213-125   Sequence 125, App
 28    140      25.4          79     15   US-10-178-213-41    Sequence 41, Appl
 29    140      25.4          82     15   US-10-178-213-83    Sequence 83, Appl
 30    139.5    25.3          76     15   US-10-178-213-281   Sequence 281, App
 31    139      25.2          47     15   US-10-178-213-84    Sequence 84, Appl
 32    138      25.0          77     15   US-10-178-213-446   Sequence 446, App
 33    138      25.0          77     15   US-10-178-213-449   Sequence 449, App
 34    138      25.0          78     15   US-10-178-213-434   Sequence 434, App
 35    137.5    24.9          77     12   US-10-072-809A-24   Sequence 24, Appl
 36    137.5    24.9          77     15   US-10-178-213-410   Sequence 410, App
 37    137      24.8          12     15   US-10-178-213-446   Sequence 446, App
 38    137      24.8           9     09   US-09-917-340-54    Sequence 54, Appl
 39    136      24.6          47     09   US-09-917-340-55    Sequence 55, Appl
 40    135.5    24.5          73     15   US-10-178-213-87    Sequence 87, Appl
 41    135      24.5          47     15   US-10-178-213-291   Sequence 291, App
 42    135      24.5          73     15   US-10-178-213-347   Sequence 347, App
 43    134      24.3          47     15   US-10-178-213-3     Sequence 3, Appli
 44    134      24.3          15     15   US-10-178-213-122   Sequence 122, App
 45    133.5    24.2          77     15   US-10-178-213-455   Sequence 455, App
```

ALIGNMENTS

RESULT 1
US-10-072-809A-18
; Sequence 18, Application US/10072809A
; Publication No. US2003021738241

GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Nicotiana alata
US-10-072-809A-18

```
Query Match       100.0%;   Score 552;  DB 12;  Length 105;
Best Local Similarity   100.0%;
Matches 105;  Conservative  0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  1 MARSLCFWAFAILARMLFVAYEVQAREENTFPGICITKPPGICITKPPCRKACISEKFTDGHCS 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 MARSLCFWAFAILARMLFVAYEVQAREENTFPGICITKPPGICITKPPCRKACISEKFTDGHCS 60

Qy 61 KILRRCLCTKPCVFDEKMKTGAEILAEEAKTLAAALLEEEIMDN 105
        |||||||||||||||||||||||||||||||||||||
Db 61 KILRRCLCTKPCVFDEKMKTGAEILAEEAKTLAAALLEEEIMDN 105
```

RESULT 2
US-10-072-809A-20
; Sequence 20, Application US/10072809A
; Publication No. US2003021738241

GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 20
; LENGTH: 105
; TYPE: PRT
; ORGANISM: peptide
US-10-072-809A-20

Query Match   95.8%; Score 529; DB 12;   Length 105;
Best Local Similarity 95.2%; Pred. No. 1.1e-52;
Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
          ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db   1 MARSLCFMAFAILAVMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60

Qy   61 KILRRCLCTKPCVFDEKMTKIGAEILAEEAKTLAAALLEEEIMDN 105
          |:|||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKPCVFDEKMIKIGAETLVEEAKTLAAALLEEEIMDN 105

RESULT 3
US-10-072-809A-51
; Sequence 51, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 51
; LENGTH: 105
; TYPE: PRT
; ORGANISM: peptide
US-10-072-809A-51

Query Match   92.0%; Score 508; DB 12;   Length 105;
Best Local Similarity 90.5%; Pred. No. 2.7e-50;
Matches 95; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
          ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db   1 MARSVCFMAFAILAVMLFVAYDVEAKDCKTESNTFPGICITKPPCRKACIEKFTDGHCS 60

Qy   61 KILRRCLCTKPCVFDEKMTKIGAEILAEEAKTLAAALLEEEIMDN 105
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKPCVFDEKVIKIGAETLAEEATTLAAALLEEEIMDN 105

RESULT 4
US-10-072-809A-52
; Sequence 52, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01

CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 52
; LENGTH: 106
; TYPE: PRT
; ORGANISM: peptide
US-10-072-809A-52

Query Match   84.9%; Score 468.5; DB 12;   Length 106;
Best Local Similarity 84.0%; Pred. No. 8.8e-46;
Matches 89; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
          ||||||||||||||||:||||||||||| |||||||||||||||||||||:|||||||||
Db   1 MARSLCFMAFAVILAVMLFVAYEVQAKSTCKAESNTFPGLCITKPPCRKACISEKFTDGKC 60

Qy   60 SKILRRCLCTKPCVFDEKMTKIGAEILAEEAKTLAAALLEEEIMDN 105
          |||||||||:|:|||||||:|||||:||||||||||||||||||:||
Db   61 SKILRRCICYKPCVFDGKMIQTGAENLAEEAKTLAAALLEEEWMDN 106

RESULT 5
US-10-072-809A-16
; Sequence 16, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 16
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Nicotiana alata
US-10-072-809A-16

Query Match   78.1%; Score 431; DB 12;   Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   26 RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPCVFDEKMTKIGAEI 85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPCVFDEKMTKIGAEI 60

Qy   86 LAEEAKTLAAALLEEEIMDN 105
          ||||||||||||||||||||
Db   61 LAEEAKTLAAALLEEEIMDN 80

RESULT 6
US-10-072-809A-14
; Sequence 14, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A.; Lay, Fung T.; Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61

```
;  SOFTWARE: Patentin version 3.0
;  SEQ ID NO 14
;    TYPE: PRT
;    LENGTH: 72
;    ORGANISM: Nicotiana alata
US-10-072-809A-14

Query Match     71.6%;  Score 395;  DB 12;  Length 72;
Best Local Similarity   100.0%;  Pred. No. 1.3e-37;
Matches   72;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS  60

Qy  61  KILRRCLCTKPC  72
        ||||||||||||
Db  61  KILRRCLCTKPC  72


RESULT 7
US-10-072-809A-50
;  Sequence 50, Application US/10072809A
;  Publication No. US20030217382A1
;  GENERAL INFORMATION:
;    APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;    TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
;    TITLE OF INVENTION: therefor
;    FILE REFERENCE: 18-01
;    CURRENT APPLICATION NUMBER: US/10/072,809A
;    CURRENT FILING DATE: 2002-09-12
;    PRIOR APPLICATION NUMBER: USSN 60/267,271
;    PRIOR FILING DATE: 2001-02-08
;    NUMBER OF SEQ ID NOS: 61
;    SOFTWARE: Patentin version 3.0
;  SEQ ID NO 50
;    LENGTH: 79
;    TYPE: PRT
;    ORGANISM: peptide
US-10-072-809A-50

Query Match     62.0%;  Score 342;  DB 12;  Length 79;
Best Local Similarity    77.5%;  Pred. No. 1.6e-31;
Matches   69;  Conservative  1;  Mismatches  9;  Indels  10;  Gaps  2;

Qy  17  LFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPCVFDE  76
        |||||||||||||||  |:||||||||||||||| ||||||||||||||||||||||
Db   1  LFVAYEVQARECARE--IFTGLCITNPQCRKACIKEXFTDGHCSKILRRCLCTKPC----  54

Qy  77  KMTKTGAEIIAEEAKTLAAALLEEEIMDN 105
           ||| ||| |||||||||||||||||||
Db  55  ---TGAETLAEEATTLAAALLEEEIMDN  79


RESULT 8
US-10-072-809A-21
;  Sequence 21, Application US/10072809A
;  Publication No. US20030217382A1
;  GENERAL INFORMATION:
;    APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;    TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
;    TITLE OF INVENTION: therefor
;    FILE REFERENCE: 18-01
;    CURRENT APPLICATION NUMBER: US/10/072,809A
;    CURRENT FILING DATE: 2002-09-12
;    PRIOR APPLICATION NUMBER: USSN 60/267,271
;    PRIOR FILING DATE: 2001-02-08
;    NUMBER OF SEQ ID NOS: 61
;    SOFTWARE: Patentin version 3.0
;  SEQ ID NO 21
;    LENGTH: 105
;    TYPE: PRT
;    ORGANISM: peptide
US-10-072-809A-21

Query Match     60.3%;  Score 333;  DB 12;  Length 105;
Best Local Similarity    62.5%;  Pred. No. 2.4e-30;
Matches   65;  Conservative  12;  Mismatches  25;  Indels  2;  Gaps  2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARB-CKTESNTFPGICITKPPCRKACISEKFTDGHC  59
        ||||:||||| | |:|||: :|| ||:|  :|| :||||:|  |:| |||:|||||:|||
Db   1  MARSIFFMAFLVLAMMLFVTVEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC  60

Qy  60  SKILRRCLCTKPCVFDEKMTKTGAEIIAEEAKTLAAALLEEEIM 103
        || :|:|| |||||||  |:| |||  ||  ||| :|:||: ||
Db  61  SKLQRKCLLCTKPCVFD-KISSEVKATLGEERAKTLSEVVLEEIM 103


RESULT 9
US-10-072-809A-8
;  Sequence 8, Application US/10072809A
;  Publication No. US20030217382A1
;  GENERAL INFORMATION:
;    APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;    TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
;    TITLE OF INVENTION: therefor
;    FILE REFERENCE: 18-01
;    CURRENT APPLICATION NUMBER: US/10/072,809A
;    CURRENT FILING DATE: 2002-09-12
;    PRIOR APPLICATION NUMBER: USSN 60/267,271
;    PRIOR FILING DATE: 2001-02-08
;    NUMBER OF SEQ ID NOS: 61
;    SOFTWARE: Patentin version 3.0
;  SEQ ID NO 8
;    LENGTH: 47
;    TYPE: PRT
;    ORGANISM: Nicotiana alata
US-10-072-809A-8

Query Match     49.6%;  Score 274;  DB 12;  Length 47;
Best Local Similarity   100.0%;  Pred. No. 5e-24;
Matches   47;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  26  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   1  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 47


RESULT 10
US-10-072-809A-25
;  Sequence 25, Application US/10072809A
;  Publication No. US20030217382A1
;  GENERAL INFORMATION:
;    APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
;    TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
;    TITLE OF INVENTION: therefor
;    FILE REFERENCE: 18-01
;    CURRENT APPLICATION NUMBER: US/10/072,809A
;    CURRENT FILING DATE: 2002-09-12
;    PRIOR APPLICATION NUMBER: USSN 60/267,271
;    PRIOR FILING DATE: 2001-02-08
;    NUMBER OF SEQ ID NOS: 61
;    SOFTWARE: Patentin version 3.0
;  SEQ ID NO 25
;    LENGTH: 47
;    TYPE: PRT
;    ORGANISM: peptide
US-10-072-809A-25

Query Match     97.9%;  Score 272;  DB 12;  Length 47;
Best Local Similarity    97.9%;  Pred. No. 8.4e-24;
Matches   46;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy  26  RBCKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   1  RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 47
```

RESULT 11
US-10-072-809A-26
; Sequence 26, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u...
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 26
; LENGTH: 47
; TYPE: PRT
; ORGANISM: peptide
US-10-072-809A-26

Query Match 33.5%; Score 185; DB 12; Length 47;
Best Local Similarity 66.7%; Pred. No. 7e-14;
Matches 30; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

```
Qy 28 CKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
Db 3  CKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHCSKLQRKCLCTKPC 47
```

RESULT 12
US-10-178-213-278
; Sequence 278, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-178-213-278

Query Match 29.5%; Score 163; DB 15; Length 78;
Best Local Similarity 41.0%; Pred. No. 4e-11;
Matches 32; Conservative 11; Mismatches 29; Indels 6; Gaps 1;

```
Qy 1  MARSLCFMAPAILARMLFVAYEV-----QARECKTESNTFPGICITKPPCRKACISEKF 54
Db 1  MASSHKFPPAVLLLLLLVTMEVAPAQAEEGRVCETDSTRFKG1CMVGTNCANICLTEGF 60

Qy 55 TDGHCSKILRRCLCTKPC 72
Db 61 TSGKCSGLRKCICTKPC 78
```

RESULT 13
US-10-178-213-290
; Sequence 290, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 290
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-178-213-290

Query Match 29.0%; Score 160; DB 15; Length 75;
Best Local Similarity 50.7%; Pred. No. 8.5e-11;
Matches 35; Conservative 6; Mismatches 24; Indels 4; Gaps 2;

```
Qy 7  FMAFAILARMLFVAYEV-----QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKIL 63
Db 8  FPAVAILL-LLVVATEVVPAQARECETESERFNGLCFVSANCAGVCNAEGFTGGKCSGLK 66

Qy 64 RRCLCTKPC 72
Db 67 RSCMCTKBC 75
```

RESULT 14
US-10-072-809A-12
; Sequence 12, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Nicotiana alata
US-10-072-809A-12

Query Match 28.4%; Score 157; DB 12; Length 33;
Best Local Similarity 100.0%; Pred. No. 7.2e-11;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 73 VFDEKMTKTGAEILAEEAKTLAAALLEEIMDN 105
Db 1  VFDEKMTKTGAEILAEEAKTLAAALLEEIMDN 33
```

RESULT 15
US-10-178-213-431
; Sequence 431, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 431
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Nicotiana benthamiana
US-10-178-213-431

  Query Match   28.1%; Score 155; DB 15; Length 78;
  Best Local Similarity   38.5%; Pred. No. 3.3e-10;
  Matches 30; Conservative 14; Mismatches 28; Indels   6; Gaps   1;

Qy   1 MARSLCFNAFAILARMLFVAYEV-----------QARECKTESNTFPGICITKPPCRKACISEKF 54
            | |: || | | | :| | |:   :||:: ||: |: | | :| |: | |
Db   1 MAKSMRFFATVLLLAMLVWATEMGPMTVAEARRCESKSQRFKGPCVRVKNCAAVCETEGF 60

Qy   55 TDGHCSKILRRCLCTKPC 72
            | : | | | : | ||
Db   61 SGGDCRGLRRRCFCTRPC 78

Search completed: January 28, 2004, 09:00:13
Job time : 372 secs

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MILYANG 23;
RA   Yun C.H., Lee M.C., Eun M.Y.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IR36;
RA   Yun C.-H., Lee J.-H., Park J.-H., Lee G.-R.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 3, PAC
RT   clone:P0043B01.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U72942; AAB17095.1; -.
DR   EMBL; AF044059; AAC00503.1; -.
DR   EMBL; AP000615; BAA85411.1; -.
DR   HSSP; P20230; 1GPT.
DR   Gramene; O24225; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
SQ   SEQUENCE   81 AA;   8757 MW;  D073B6ED5ACGBA27 CRC64;

Query Match      26.7%;  Score 147.5;  DB 10;  Length 81;
Best Local Similarity   37.0%;  Pred. No.2;1e-10;
Matches   27;  Conservative  13;  Mismatches  26;  Indels   7;  Gaps    1;

Qy          7 FMAFAILARMLFVAYEV---------QAESCKTESNTFPGICITKPPCRKACISEKFTDGHC 59
               ||::| ::|            ::|||:||::|::|:|||||:||:.| .||:|:
Db          8 FSAMLLMVLLAATGEMGGPVWVAEARTCESQSHRFKGPCARKANCASVCNTEGFPDGYC 67

Qy         60 SKILRRCLCTKPC 72
               :|||:|:|||||
Db         68 HGVRRRCMCTKPC 80
```

Search completed: January 28, 2004, 08:48:08
Job time : 37 secs

```
RESULT 13
ID   Q9C947
AC   Q9C947;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 6.0 kDa protein.
GN   T7P1.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
DR   EMBL; AC018908; AAG51654.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   55 AA;   5977 MW;   167C3B51F9BF0F52 CRC64;

Q9C947    PRELIMINARY;    PRT;    55 AA.

Query Match
Best Local Similarity   50.0%;   Pred. No. 9.2e-11;
Matches  25;  Conservative   7;  Mismatches  18;  Indels  0;  Gaps  0;

Qy  23 VQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
        : |||.|||.||.|.|.|||.||:|.|| ||.:|:: |:|||||||||
Db   6 VEARTCETSSNLFNGPCLSSSNCANVCHNEGFSDGDCRGFRRRCLCTRPC 55
```

```
RESULT 14
ID   Q9MB66
AC   Q9MB66;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Thionin like protein.
GN   NT-THIONIN.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takemoto D., Kawakita K.;
RT   "Molecular cloning of fungal elicitor inducible genes of Tobacco.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB034956; BAA95697.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
SQ   SEQUENCE   78 AA;   8577 MW;   1993CA5C94C8AE51 CRC64;

Q9MB66    PRELIMINARY;    PRT;    78 AA.

Query Match
Best Local Similarity   37.2%;   Pred. No. 1.3e-10;
Matches  29;  Conservative  13;  Mismatches  30;  Indels  2;  Gaps  1;

Qy   1 MARSLCFMAFAILARMLFVAYEV----------QARECKTESNTFPGICITKPPCRKACISEKF 54
         ||.:.:.|:.||.||:|..| ||   |   ||.||:|:.:|.|.|.:|..|
Db   1 MANSMRFFATVLLIALLVTA-EMGPMTIAEARTCESQSHRFKGPCSRDSNCATVCLTEGF 60

Qy  55 TDGHCSKILRRCLCTKPC 72
         :.| || |||||:||.|
Db  61 SGGDCRGFRRRCFCTRPC 78
```

```
RESULT 15
ID   Q24225
AC   Q24225;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Proteinase inhibitor.
GN   RPI OR RGPI9.
OS   Oryza sativa (Rice).
```

```
Q024225    PRELIMINARY;    PRT;    81 AA.

Query Match
Best Local Similarity   37.0%;   Pred. No. 1.3e-10;
Matches  29;  Conservative  13;  Mismatches  30;  Indels  2;  Gaps  1;

Qy   1 MARSLCFMAFAILARMLFVAYEV----------QARECKTESNTFPGICITKPPCRKACISEKF 54
Db   1 MANSMRFFATVLLIALLVTA-EMGPMTIAEARTCESQSHRFKGPCSRDSNCATVCLTEGF 60

Qy  55 TDGHCSKILRRCLCTKPC 72
Db  61 SGGDCRGFRRRCFCTRPC 78
```

```
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Protease inhibitor.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Essex; TISSUE=leaf;
RA   Maitra N., Cushman J.C.;
RT   "Characterization of a Drought-Induced Soybean cDNA Encoding a Plant
RT   Defensin (Accession No. U12150) (PGR98-213).";
RL   Plant Physiol. 118:1536-1536(1998).
DR   EMBL; U12150; AAC97524.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002118; Gamma-thionin.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
SQ   SEQUENCE   79 AA;   8772 MW;   24B241BFF67058E3 CRC64;

Query Match
Best Local Similarity   36.7%;
Matches  29;  Conservative  16;  Mismatches  27;  Indels  7;  Gaps  1;

Qy   1 MARSLCFMAFAILARMLFVAYEV----------QARECKTESNTFPGICITKPPCRKACISEK 53
Db   1 MSRSVPLVSTICVLLLLLVATEMMGPTMVAEARTCESQSHRFKGPCLSDTNCGSVCRTER 60

Qy  54 FTDGHCSKILRRCLCTKPC 72
Db  61 FTGHCRGFRRRCFCTKXC 79
```

RESULT 9
```
ID   Q9SEM1      PRELIMINARY;    PRT;    84 AA.
AC   Q9SEM1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Gamma-thionin 1 precursor.
GN   THION1.
OS   Capsicum annuum (bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Hanbyul; TISSUE=Leaf;
RA   Hwang B.K., Lee S.C., Kim Y.J., Hong J.K.;
RT   "Molecular cloning and pathogen-induced expression of a thionin gene
RT   in pepper plants.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF112869; AAF18936.1; -.
DR   InterPro; IPR002118; Knot1.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL        1     21
FT   CHAIN        22     84       GAMMA-THIONIN 1.
SQ   SEQUENCE   84 AA;  9470 MW;  03B211731C04E493 CRC64;

Query Match             35.2%;  Score 194.5;  DB 10;  Length 84;
Best Local Similarity   44.4%;  Pred. No. 3.7e-16;
Matches   44;  Conservative   10;  Mismatches   28;  Indels   17;  Gaps   2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACI-SEKFTDGHC  59
            ||| || ||| || |||  |    :|   |   | ||    |    |  :| |||||
Db   1  MARSIYFMAFLVLLAMTVFVAYGVGVQGKEICCKELIKPVKSSSDP:CQKLCMEKKYEDGHC  60

Qy  60  SKILRRCLCTKPCVFDERMTKTGAEILAEBAKTLAAALL  98
            |||  |||      |              ||
Db  61  FTILSKCLCMKRC------------------NAKTLATELL  83
```

RESULT 10
```
ID   Q8W4V6      PRELIMINARY;    PRT;    78 AA.
AC   Q8W4V6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Proteinase inhibitor precursor.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Hanbyul; TISSUE=Leaf;
RA   Hwang B.K., Do H.M.;
RT   "Differential expression of pepper proteinase inhibitor gene in
RT   response to pathogen infection, abiotic elicitors and osmotic stresses
RT   in Capsicum annuum";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF442388; AAL35366.1; -.
DR   InterPro; IPR002118; Knot1.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.

Query Match             29.9%;  Score 165;  DB 10;  Length 78;
Best Local Similarity   39.7%;  Pred. No. 1.4e-12;
Matches   31;  Conservative   14;  Mismatches   27;  Indels   6;  Gaps   1;

Qy   1  MARSLCFMAFAILARMLFVAYE------VQARECKTESNTFPGICITKPPCRKACISEKF  54
            || | |   | ||  |         |   |  :   ||  | :   |  |
Db   1  MAHSMRFFAIVLLLAVLVMATEMGPMRIVEARTCESQSHRFKGVCASETNCASVCQTEGF  60

Qy  55  TDGHCSKILRRCLCTKPC  72
             | || |  ||| | ||
Db  61  SGGDCRGFRRRCFCTRPC  78
```

RESULT 11
```
ID   Q8LEG6      PRELIMINARY;    PRT;    76 AA.
AC   Q8LEG6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protease inhibitor II.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY085425; AAM62652.1; -.
DR   InterPro; IPR002118; Knot1.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Signal.
FT   SIGNAL        1     20
SQ   SEQUENCE   78 AA;  8689 MW;  E10E218F5A6CB5BE CRC64;

Query Match             28.5%;  Score 157.5;  DB 10;  Length 76;
Best Local Similarity   42.9%;  Pred. No. 1.2e-11;
Matches   30;  Conservative   11;  Mismatches   28;  Indels   1;  Gaps   1;

Qy   3  RSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKI  62
            | |   | |  |            |  : | | :     |        | : | :||
Db   8  KSALLLIFMILAIVM-GPVRVEARTCETSSNLFNGPCLSSSNCANVCHNEGFRSDGDCRGF  66

Qy  63  LRRCLCTKPC  72
            ||| ||:||
Db  67  RRRCLCTRPC  76
```

RESULT 12
```
ID   Q39807      PRELIMINARY;    PRT;    79 AA.
AC   Q39807;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
```

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Old Glory Blue;
RA   Lay F.T., Brugliera F., Anderson M.A.;
RT   "Isolation and properties of floral defensins from Nicotiana alata and
RT   Petunia hybrida.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF507976; AAN64751.1; -.
SQ   SEQUENCE  101 AA;  11049 MW;  8E5AFE2BD4052D0B CRC64;

Query Match     58.1%;  Score 320.5;  DB 10;  Length 101;
Best Local Similarity   60.4%;
Matches  64;  Conservative  10;  Mismatches  25;  Indels   7;  Gaps   2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACIS--EKFTDGH   58
        ||||.||.|.|||.|||||||.....|.|..:.|.|:|||.|..|:.|:|   |||||||
Db   1  MARSICFFAVAILALMLFAAYETEAGTCKAECPTWEGICINKAPCVKCCKRAQPEKFTDGH  60

Qy  59  CSKILRRCLCTKPCA-----TEEATATIANEVKTMAEALVEEDMME  101
        |||||||||||||||     ||||||||||||||||||||||||||
Db  61  CSKILRRCLCTKPCA-----TEEATATIANEVKTMAEALVEEDMME  101

RESULT 6
ID   Q8H6Q1        PRELIMINARY;      PRT;   103 AA.
AC   Q8H6Q1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral defensin-like protein 1.
GN   D1.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Old Glory Blue;
RA   Lay F.T., Brugliera F., Anderson M.A.;
RT   "Isolation and properties of floral defensins from Nicotiana alata and
RT   Petunia hybrida.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF507975; AAN64750.1; -.
SQ   SEQUENCE  103 AA;  11361 MW;  A8B629A3E006A0D01 CRC64;

Query Match     54.8%;  Score 302.5;  DB 10;  Length 103;
Best Local Similarity   57.3%;  Pred. No. 2.4e-29;
Conservative  10;  Mismatches  33;  Indels   1;  Gaps   1;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS  60
        ||||.||.|.|||.|||||||......|.|..:.|.|.:..:.|:..|:..::||.||||
Db   1  MARSICFFAVAILALMLFAAYDAEAATCKAECPTWDSVCINKKPCVACCKKAKFSDGHCS  60

Qy  61  KILRCLCTKPCVFPDEKMKTGAEILAEEAKTLIAAALLEEEIM  103
        |||.|||||.||| :||:|.:|||||:.|:.:||.|||| ::|
Db  61  KILRRCLCTKBCVF-EKTEATQTEFTKDVNTLAEALLEADMM  102

RESULT 7
ID   Q9XHE3        PRELIMINARY;      PRT;   107 AA.
AC   Q9XHE3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Putative gamma-thionin.
OS   Capsicum chinense.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=80379;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. habanero;
RX   MEDLINE=99372458; PubMed=10447467;
RA   Aluru M., Curry J., O'Connell M.A.;
RT   "Nucleotide sequence of a defensin or gamma-thionin-like gene
RT   (Accession No. AF128239) from habanero chile (PGR 99-070).";
RL   Plant Physiol; 120:633-633(1999).
DR   EMBL; AF128239; AAD21200.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE  107 AA;  11964 MW;  B173320A754284C4 CRC64;

Query Match     51.4%;  Score 284;  DB 10;  Length 107;
Best Local Similarity   57.5%;  Pred. No. 4.8e-27;
Matches  61;  Conservative   9;  Mismatches  32;  Indels   4;  Gaps   3;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARE--CKTESNTFPGICITKPPCRKACISE-KFTDG   57
        ||||..|||.|.||.|||||...:|.    |||.|.:||.|...|:.||:..|: .|.||
Db   1  MARSIYFMAFLVLAVTLFVANGVQGQONNICKTTSKHFEKGLCFADSKCRKVCIQSDKFEDG  60

Qy  58  HCSKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLIAAALLEEEIM  103
        |||||:.|||||.||||.:   :.:|...:.|.|||..|.|:|||:|
Db  61  HCSKLQRKCLCTKNCVFDNIPNDVGT--IIVQDAKTLEAQLIEEEIL  105

RESULT 8
ID   Q9SEM4        PRELIMINARY;      PRT;    84 AA.
AC   Q9SEM4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Thionin-like protein.
GN   PEPTHI.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20064969; PubMed=10598099;
RA   Oh B.J., Ko M.K., Kostenyuk I., Shin B., Kim K.S.;
RT   "Coexpression of a defensin gene and a thionin-like gene via different
RT   signal transduction pathways in pepper and Colletotrichum
RT   gloeosporioides interactions.";
RL   Plant Mol. Biol. 41:313-319(1999).
DR   EMBL; AF112443; AAF16413.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   84 AA;   9484 MW;  8182116E39E4F166 CRC64;

Query Match     35.8%;  Score 197.5;  DB 10;  Length 84;
Best Local Similarity   45.5%;  Pred. No. 1.6e-16;
Matches  45;  Conservative   9;  Mismatches  28;  Indels  17;  Gaps   2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACI-SEKFTDGHC   59
        ||||..|||.|.||..||||||.|.:|  |.|.    :.|.|:|||.|.   :|:|.||||
Db   1  MARSIYFMAFLVLAMTLFVAYGVQGKEICCKELTKPVKCSSDPLCQKLCMEKEKYEDGHC  60

Qy  60  SKILRRCLCTKPCVFDEKMKTKTGAEILAEEAKTLAAALL  98
        ..|:..||||.:|                  |:||:| ||
Db  61  FTILSKCLCMKRC---------------NAKTLATELL  83
```

```
O24105
ID   O24105               PRELIMINARY;      PRT;   105 AA.
AC   O24105;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Gamma-thionin.
GN   NETHIO2.
OS   Nicotiana excelsior.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=61185;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada S.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   McKendree W.L., Doostdar H., McCollum T.G., Mayer R.T.;
RT   "cDNA cloning and expression of a gene (Accession No. 297064) from
RT   Citrus paradisi roots similar to bacterial YRN1 and NEAHIO proteins
RT   and an mRNA from Brassica oleracea that is wound and dark inducible
RT   (PGR97-127).";
RL   Plant Physiol. 115:314+314(1997).
DR   EMBL; AB005260; BAA21114.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
SQ   SEQUENCE   105 AA;  11635 MW;  B7C586C8DB565DF CRC64;
```

```
Query Match              92.0%;  Score 508;  DB 10;  Length 105;
Best Local Similarity    90.5%;  Pred. No. 1.4e-54;
Matches     95;  Conservative     5;  Mismatches     5;  Indels     0;  Gaps     0;

Qy          1 MARSLCFMAFAILARMLFVAYEVQAREKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARSVCFMAFAILAVMLFVAYDVEAKDCKTESNTFPGICITKPPCRKACIKEKFTDGHCS 60.

Qy         61 KILRRCLCTKPCVFDEKMTKTGAEIIAERAKTLAAALLEEEIMDN 105
              |||||||||||||||||||| |||||||||||||||||||||||
Db         61 KILRRCLCTKPCVFDEKMIKTGAETLAEEATTLAAALLEEEIMDN 105
```

```
RESULT 3
O24104
ID   O24104               PRELIMINARY;      PRT;   79 AA.
AC   O24104;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Gamma-thionin (Fragment).
GN   NETHIO1.
OS   Nicotiana excelsior.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=61185;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada S., Komori T., Imaseki H.;
RT   "cDNA cloning of gamma-thionin from Nicotiana excelsior.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB005265; BAA21113.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
```

```
FT   NON_TER          1        1
SQ   SEQUENCE   79 AA;  8785 MW;  FB507E957D78D285 CRC64;
```

```
Query Match              62.0%;  Score 342;  DB 10;  Length 79;
Best Local Similarity    77.5%;  Pred. No. 2.6e-34;
Matches     69;  Conservative     1;  Mismatches     9;  Indels    10;  Gaps     2;

Qy         17 LFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPCVFDE 76
              |||||||||||: |||:||||||||||||||||||||||||||||||||||||||||||
Db          1 LFVAYEVQARECARE--IFTGLCITNPQCRKACIKEKFTDGHCSKILRRCLCTKPC---- 54

Qy         77 KMTKTGAEIIAEBAKTLAAALLEEEIMDN 105
              ||| |||||||||||||||||||||
Db         55 ----TGAETLAEEATTLAAALLEEEIMDN 79
```

```
RESULT 4
Q40128
ID   Q40128               PRELIMINARY;      PRT;   105 AA.
AC   Q40128;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Flower-specific gamma-thionin-like protein/acidic protein
DE   precursor.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF36; TISSUE=Pistil;
RX   MEDLINE=95375233; PubMed=7647301;
RA   Milligan S.B., Gasser C.S.;
RT   "Nature and regulation of pistil-expressed genes in tomato.";
RL   Plant Mol. Biol. 28:691-711(1995).
DR   EMBL; U20591; AAA80496.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL           1       26
FT   CHAIN           27       73       GAMMA-THIONIN-LIKE PROTEIN.
FT   CHAIN           74      105       ACIDIC PROTEIN.
SQ   SEQUENCE   105 AA;  11914 MW;  ADC987ECB620E814 CRC64;
```

```
Query Match              60.3%;  Score 333;  DB 10;  Length 105;
Best Local Similarity    62.5%;  Pred. No. 4.5e-33;
Matches     65;  Conservative    12;  Mismatches    25;  Indels     2;  Gaps     2;

Qy          1 MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
              ||||:|||||||||| ||:||||||:||  ||||||||||||||||||||||||||||
Db          1 MARSIFMAFLVLAMVLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC 60

Qy         60 SKILRRCLCTKPCVFDEKMTKTGAEIIAEBAKTLAAALLEEEIM 103
              ||||||||||||||:| |||  :|:||||||||||||||||||
Db         61 SKLQRKCLCTKPCVFD-KISSSVKATLGEEAKTLSEVVLEBEIM 103
```

```
RESULT 5
Q8H6Q0
ID   Q8H6Q0               PRELIMINARY;      PRT;   101 AA.
AC   Q8H6Q0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral defensin-like protein 2.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OM protein - protein search, using sw model

                              GenCore version 5.1.6
                              Copyright (c) 1993 - 2004. Compugen Ltd.

Run on:              January 28, 2004, 08:46:51 ; Search time 35 seconds
                              (without alignments)
                              774.158 Million cell updates/sec

Title:               US-10-072-809B-18
Perfect score:       552
Sequence:            1 MARSLCFWAFAILLARMLFVA..........LAEEAKTLAAALLEEEIMDN 105

Scoring.table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched:            830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters:        830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
          1: sp_archea:*
          2: sp_bacteria:*
          3: sp_fungi:*
          4: sp_human:*
          5: sp_invertebrate:*
          6: sp_mhc:*
          7: sp_mammal:*
          8: sp_organelle:*
          9: sp_phage:*
         10: sp_plant:*
         11: sp_rodent:*
         12: sp_virus:*
         13: sp_vertebrate:*
         14: sp_unclassified:*
         15: sp_rvirus:*
         16: sp_bacteriap:*
         17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result            %      Query
No.   Score  Match  Length  DB   ID        Description
----  -----  -----  ------  --  --------   -----------
  1    546   98.9    105    10  Q8GTMO     q8gtmo nicotiana a
  2    508   92.0    105    10  Q8GTMO     q8gtmo nicotiana a
  3    342   62.0     79    10  Q24105     q24105 nicotiana e
  4    333   60.3    105    10  Q24104     q24104 nicotiana e
  5    320.5  58.1   101    10  Q40128     q40128 lycopersico
  6    302.5  54.8   101    10  Q8H6Q0     q8h6q0 petunia hyb
  7    284    51.4   107    10  Q8H6Q1     q8h6q1 petunia hyb
  8    197.5  35.8    84    10  Q9XHE3     q9xhe3 capsicum ch
  9    194.5  35.2    84    10  Q9SEM4     q9sem4 capsicum an
 10    165    29.9    78    10  Q8W4V6     q8w4v6 capsicum an
 11    157.5  28.5    76    10  Q9SEM1     q9sem1 capsicum an
 12    150.5  27.3    79    10  Q8LEG6     q8leg6 arabidopsis
 13    149    27.0    55    10  Q39807     q39807 glycine max
 14    149    27.0    78    10  Q9MB66     q9mb66 arabidopsis
 15    147.5  26.7    81    10  Q24225     q24225 oryza sativ
 16    142.5  25.8    87    10  Q48T3      q48t3 pyrus pyrif
```

```
17   136   24.6   72   10   Q9XG53    lycopersico
18   136   24.6   78   10   Q945D8    castanea sa
19   135   24.5   81   10   Q948T2    pyrus pyrif
20   134   24.5   77   10   P82788    arabidopsis
21   131   23.7   87   10   Q948T4    pyrus pyrif
22   129   23.4   49   10   Q93WS9    musa acumin
23   128.5  23.3   73   10   Q93WS9    phaseolus c
24   125.5  22.7   77   10   Q9FUP3    phaseolus c
25   122   22.1   73   10   Q39403    brassica oa
26   122   22.1   73   10   Q9FUP3    brassica ca
27   122   22.1  113   10   Q9FQ14    citrus para
28   119   21.6   82   10   Q9FG14    citrus para
29   118   21.4   77   10   Q942U6    oryza sativ
30   118   21.4   83   10   Q81698    triticum ae
117.5 21.3   75   10   Q81698    triticum ae
31   116   21.0   83   10   Q8H766    elaeis guin
32   112   20.3   73   10   Q8GTL2    picea abies
33   112   20.3   80   10   Q39894    glycine max
34   112   20.3   80   10   Q8GTL2    glycine max
35   106   19.2   80   10   Q40779    picea abies
36   106   19.2   80   10   Q9FFP8    arabidopsis
37   104   18.8   80   10   Q9FI22    arabidopsis
38   103   18.7   42   10   Q9FI22    arabidopsis
39   103   18.7   80   10   Q9FZ31    arabidopsis
40   101   18.3   41   10   Q94IN7    eutrema was
41   100   18.1   75   10   Q9FS38    brassica ol
41   100   18.1   71   10   Q94IN7    helianthus
41    99   17.9    5   10   Q8h6k0    brassica ra
42    97   17.9   71   10   Q8IRD7    helianthus
42    97   17.6   71   10   Q8IRD7    drosophila
43    95   17.6   91   10   Q8W434    vigna radia
43    95   17.6   91   10   Q9AY29    zea mays (m
44    95   17.2   91   10   Q9AY30    zea mays (m
44    95   17.2   91   10   Q9AY28    zea mays (m
45   94.5   17.1   82   10   Q39999    hordeum vul
```

ALIGNMENTS

RESULT 1
Q8GTMO
ID   Q8GTMO;   PRELIMINARY;   PRT;   105 AA.
AC   Q8GTMO;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Flower-specific defensin precursor.
GN   NAD1.
OS   Nicotiana alata (Winged tobacco) (Persian tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lay F.T., Schirra H.J., Scanlon M.J., Anderson M.A., Craik D.J.;
RT   "The three-dimensional solution structure of NaD1, a new floral
RT   defensin from Nicotiana alata and its application to a homology model
RT   of the crop defense protein alfAFP.";
RL   J. Mol. Biol. 0:0-0(2003).
DR   EMBL; AF509566; AAN70999.1; -.
OX   NCBI_TaxID=4087;
CHAIN   105 AA;   26   72   FLOWER-SPECIFIC DEFENSIN.
SQ   SEQUENCE   105 AA;   11722 MW;   DA7F41736CEE6AC3 CRC64;

Query Match   98.9%;   Score 546;   DB 10;   Length 105;
Best Local Similarity   99.0%;   Pred. No. 3.1e-59;
Matches 104; Conservative   0; Mismatches   1;   Indels   0; Gaps   0;

Qy   1 MARSLCFWAFAILLARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MARSLCFWAFAILLAMMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60

Qy   61 KILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIMDN 105
           |||||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIMDN 105

RESULT 2

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 28, 2004, 08:44:46

Title: US-10-072-809B-17
Perfect score: 541
Sequence: 1 atggctcgctcccttgtgctt..........tggttaaaaaaaaaaaa 541
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Post-processing: Minimum Match 0%
Maximum Match 100%
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database : N_Geneseq_19Jun03:*
Listing first 45 summaries

Scoring: ; Search time 219 Seconds
(without alignments)
6668.475 Million cell updates/sec

```
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	490.8	90.7	566	20	AAV70142	Nicotiana excelsio
2	377.6	69.9	558	20	AAV70143	Nicotiana panicula
3	362.8	67.1	456	22	AAV70141	Nicotiana excelsio
4	189.8	35.1	564	12	AAQ10262	Ovary tissue trans
5	189.8	35.1	564	14	AAQ34940	p2130 coding seque
6	189.8	35.1	564	17	AAT31823	Tomato p2130 cDNA
7	189.8	35.1	564	18	AAT48813	cDNA clone p2130 c
8	189.8	35.1	564	24	AB221970	Plasmid p2130 cDNA

ALIGNMENTS

RESULT 1
AAV70142
ID   AAV70142 standard; cDNA to mRNA; 566 BP.
XX
AC   AAV70142;
XX
DT   03-FEB-1999 (first entry)
XX
DE   Nicotiana excelsior thionine gene NeTHI2.
XX
KW   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
KW   NpTHI1; salt stress; resistance; ds.
XX
OS   Nicotiana excelsior.
XX
PN   JP1029538O-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997; 97JP-0120179.
PR   23-APR-1997; 97JP-0120179.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
PT
FH   Key   Location/Qualifiers
FT   CDS   33..350
FT   /*tag= a
XX
DR   WPI; 1999-038278/04.
DR   P-PSDB; AAW83132.

Tomato p2130 genom
Calgene Lambda 140
Plasmid p2130 cDNA
pZ130 contg. Calge
pZ130 cDNA
Calgene lambda 140
Calgene lambda 140
Pepper defensin pr
DNA encoding Capsi
Haematopoietic cel
Haematopoietic cel
Chemically treated
Human immune syste
Corn tassel-derive
Human angiogenesis
Haematopoietic cel
Haematopoietic cel
Human immune syste
Corn tassel-derive
Nucleic acid seque
Buchnera sp. genom
Chemically treated
Human metastasis a
Human chemically m
Signal transductio
Chemically treated
Human gene regulat
Human prostate exp
Arabidopsis thalia
Chemically treated
Human gene regulat
Tumour suppressor
Chemically pretrea
Human immune syste
DNA transcription
Human immune syste

```
 9  147.4   27.2  3528  17 AAT31824
10  147.4   27.2  3528  18 AAT48816
11  147.4   27.2  3528  24 AB221971
12  147.4   27.2  4383  12 AAQ10263
13  147.4   27.2  4383  13 AAQ10319
14  147.4   27.2  4383  14 AAQ35143
15  147.4   27.2  4383  14 AAQ34941
16  95.6    17.7   506  22 AAF83907
17  95.6    17.7   548  24 ABX12800
18  47.4    8.8   19634  25 ABZ10016
19  47.4    8.8   19634  25 ABZ10162
20  45.2    8.4   34548  25 ABL33712
21  44.6    8.2   13326  24 ABL37712
22  43.4    8.0    283   24 ABZ10246
23  43.4    8.0   5611   25 ABQ67070
24  43.2    8.0   8056   25 ABZ10246
25  42.4    7.8   1501   25 ABZ10188
26  42.4    7.8   6042   24 ABL33592
27  42.4    7.8   8056   25 ABZ10100
28  42      7.8    297   24 ABL75367
29  42      7.8   1692   20 AAX99563
30  42      7.8  640681  24 ABA92787
31  41.8    7.7   6137   24 ABL70128
32  41.8    7.7   6137   24 ABL34451
33  41.8    7.7   6250   24 ABN00214
34  41.8    7.7  13133   24 ABK21230
35  41.8    7.7  18283   24 ABL70502
36  41.8    7.7  18283   24 AAS61363
37  41.6    7.7   506    23 ABVS8067
38  41.6    7.7   1438   24 ABZ16790
39  41.6    7.7   6754   24 ABL70346
40  41.4    7.7   6754   24 AAS61305
41  41.4    7.7  17280   22 AAS46772
42  41.4    7.6   6665   24 AAS45298
43  41      7.6   6665   24 ABL32082
44  41      7.6   6665   24 ABK28129
45  41      7.6  10433   24 ABL32378
```

```
XX
PT   Thionine gene derived by salt stress - used to deliver improved salt
PT   stress to plants
XX
PS   Claim 3; Page 5; 6pp; Japanese.
XX
CC   The present sequence represents a thionine gene from Nicotiana excelsior
CC   derived NeTHI2. The thionine protein has an effect of improving the salt
CC   stress resistance of a plant. The gene can improve the salt stress
CC   resistance of a plant.
XX
SQ   Sequence 566 BP; 178 A; 94 C; 108 G; 186 T; 0 other;

Query Match     90.7%;  Score 490.8;  DB 20;  Length 566;
Best Local Similarity   94.9%;
Matches  507;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy        1 ATGGCTCGCTCCTTGTCGCTTCATGGCCATTTGCTATCTTGGCCAAGGATGCTCTTTGTTGCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       33 ATGGCTCGCTCCGTGGCTTCATGGCCATTTGCTATCTTGGCAGTCATGCTCCTTTGTTGCC 92

Qy       61 TATGAGGTGCAAGCTAGAGAAATGCAAAACAGAAAGCAACACATTCCTGGAATATGCATT 120
Db       93 TATGATGTGGAAGCTAAAGATTGCAAAACAGAAAGCAATACACATTCCCTGGAATATGCAT 152

Qy      121 ACCAAACCACCATGCAGAAAAGCTTGTATCACAGTGATGAGAAATTACTGATGTCATTGTAGC 180
Db      153 ACCAAACCACCATGCAGAAAAGCTTGTATCAAAGAGAAATTACTGATGTCATTGTAGC 212

Qy      181 AAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGAATGAGAAGATGATCAAA 240
Db      213 AAAAATCCTCAGAAGGTCTCTATGCACTAAGCCATGTGTTTGATGAGAAGATGATCAAA 272

Qy      241 ACAGAGAGCTGAAAATTTTGGCTGCGAGAAGCAAAAACTTGGCTGCAGCTTTGCTTGAAGAA 300
Db      273 ACAGGAGCTGAAAACTTTAGTGTGAGAAGCAACAACTTTGGCTGCAGCTTTGCCTGAAGAA 332

Qy      301 GAGATAAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGATGCAGTATCACACATAAT 360
Db      333 GAGATAAATGGATAAACTAATTAGAAGATTAGAAGAGTCACACATAAT 392

Qy      361 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTTAAATGGCTTTTTAGTAGC 420
Db      393 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTTAAATTGGCTTTTTAGTAGC 452

Qy      421 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAATCCTTGTGCAAATCCTTGCAC 480
Db      453 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAATCCTTGTGCAAATCCTTGCACTAAGT 512

Qy      481 TTATTTGTGTACTTTTAATGAAAAATGAACCTTCTATGGTTCTTCTTTGGTTAAAAAAAA 534
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      513 TTATTCGTGTACTTTTAATGAAAATCACCTTCTATGGTTTTTCTTTGGTTTAAAAAAAA 566
```

```
PN   JP10295380-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997;  97JP-0120179.
XX
PR   23-APR-1997;  97JP-0120179.
XX
PA   (NISB ) JAPAN TOBACCO INC.
XX
CC   The present sequence represents a thionine gene from Nicotiana
CC   paniculata derived NpTHI1. The thionine protein has an effect of
CC   improving the salt stress resistance of a plant. The gene can improve
CC   the salt stress resistance of a plant.
XX
DR   WPI; 1999-038278/04.
DR   P-PSDB; AAW83133.
XX
SQ   Sequence 558 BP; 175 A; 87 C; 112 G; 184 T; 0 other;

Query Match     69.8%;  Score 377.6;  DB 20;  Length 558;
Best Local Similarity   86.2%;
Matches  464;  Conservative    0;  Mismatches   44;  Indels   30;  Gaps    3;

Qy        1 ATGGCTCGCTCCTTGTCGCTTCATGGCCATTTGCTATCTTGGCCAAGGATGCTCTTTGTTGCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       48 ATGGCTCGCTCCTTGTGTGCTTCATGGCCATTTGCTATCTTGGCAGTCATGCTCCTTTGTTGCC 107

Qy       61 TATGAGGTGCAAGCT---AGAGAAATGCAAAACAGAAAGCAACACATTCCTGGAATATGC 117
Db      108 TATGAGGTTGCAAGCTTAAGAGTACTTGCAAAGCAATACACATTCCCTGGAATTATGC 167

Qy      118 ATTACCAAACCACCATGCAGAAAAAGCTTGTATCAGTGATGAGAAATTACTGATGTCATTGT 177
Db      168 ATTACCAAAACCACACATGCAGAAAAAGCTTGTCTCCAGTGAGAAAATTACTGATGTCAAAGT 227

Qy      178 AGCAAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGATC 237
Db      228 AGCAAAAATCCTCAGAAGGTGCATTTGCTACAAGCCATGTGTATTTGATGGAAGAAGATATC 287

Qy      238 AAAACAGAGAGCTGAAAATTTTGGCTGCGAGAAGCAAAAACTTGGCTGCAGCTTTGCTGAA 297
Db      288 CAAACAGGAGCTGAAAATTTTGGCCGAGAAGCAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAA 347

Qy      298 GAAGAGATAAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGATGCAGTATCACACAT 357
Db      348 GAAGAGATGAATGGATAACTAATTAGAGATTAGAGAATTAAAGGATGAAGTGTCACACAT 407

Qy      358 AATAAAAGTTTCTACCTTTCTTAAAAGTGTAGCTTAAATGGCTTTTTAGT 417
Db      408 AATAAAG-TGCTGGCCTTTCTTAAAAGTGTAGCTTAAATGTTGGCTTTTTAGT 466

Qy      418 AGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAATCCTTGTGCAAATCCTTGCACTA 477
Db      467 AGCCGTTTGTTTGTTACACTTTAAATAAGTGTGGCACATCAATCCTTTGT----------- 512

Qy      478 AGTTTATTTGTGTACTTTTAATGAAAAATGAACCTTCTATGGTTCTTCTTTTAAAAAAAAA 535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      513 -------------TACTTTTAATGAAAATGAACCTTCTATGGTTTTTCTTTGGTTTAAAAAAAA 558
```

```
RESULT 2
AAV70143
ID   AAV70143 standard; cDNA to mRNA; 558 BP.
XX
AC   AAV70143;
XX
DT   03-FEB-1999  (first entry)
XX
DE   Nicotiana paniculata thionine gene NpTHI1.
XX
KW   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
KW   NpTHI1; salt stress; resistance; ds.
XX
OS   Nicotiana paniculata.
XX
```

```
RESULT 3
AAV70141
ID   AAV70141 standard; cDNA to mRNA; 456 BP.
XX
AC   AAV70141;
XX
DT   03-FEB-1999  (first entry)
```

```
FH   Key             Location/Qualifiers
FT   CDS             48..368
FT                   /*tag= a
XX
```

```
XX
DE   Nicotiana excelsior thionine gene NeTHI1.
XX
KW   NpTHI1; salt stress; resistance; ds.
KW   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
XX
OS   Nicotiana excelsior.
XX
PN   JP1029S380-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997;  97JP-0120179.
XX
PR   23-APR-1997;  97JP-0120179.
XX
PA   (NISB ) JAPAN TOBACCO INC.
XX
DR   WPI; 1999-038278/04.
DR   P-PSDB; AAW83131.
XX
PT   Thionine gene derived by salt stress - used to deliver improved salt
PT   stress to plants
XX
PS   Claim 1; Page 4; 6pp; Japanese.
XX
CC   The present sequence represents a thionine gene from Nicotiana excelsior
CC   derived NeTHI1. The thionine protein has an effect of improving the salt
CC   stress resistance of a plant. The gene can improve the salt stress
CC   resistance of a plant.
XX
SQ   Sequence 456 BP; 149 A; 73 C; 88 G; 146 T; 0 other;
```

```
FH   Key             Location/Qualifiers
FT   CDS             1..240
FT                   /*tag= a
```

Query Match   67.1%; Score 362.8; DB 20; Length 456;
Best Local Similarity 88.3%; Pred. No. 3.2e-79;
Matches 429; Conservative   0; Mismatches 27; Indels 30; Gaps 2;

```
Qy     49 CTCTTTGTTGCCTATGAGGTGCCAAGCTGTAGACAAGTGCAAAAAGAGAAAAGCACACATTTCCT 108
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCTTTGTTGCCCATGAGGTGCAAGCTGTAGACAAGTGCAAGAGAAATTTCA-------CT 54

Qy    109 GGAATATGCAATTACCAAACCACCATGCGAGAAAAGCTTGTATCAGTGCAGAAAATTTACTGAT 168
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     55 GGACTAGCAATTACCAATCCACCAATGCAGAAAAGCTTGTATCAAAGAGAAATTTACTGAT 114

Qy    169 GGTCATTGTAGCAAAATCCTCAGAAGGTGCCTTATGTACTAAGCCATGTGTGTTTGATGAG 228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    115 GGTCATTGTAGCAAAATCCTCAGAAGGTGTCTATGCACTAAGCCATG----------- 161

Qy    229 AAGATGACTAAAACAGGAGCTGAAATTTGGGCTGAGGAAGCAAAAACTTTGGCTGCAGCT 288
                                          ||||||||||||||||||||||||||||||
Db    162 -----------CACAGGAGCTGAAACTTTAGCTGCTGAAGACAACAACTTTGGCTGCAGCT 210

Qy    289 TTGCTTGAAGAAGAGATAAATGGATAAACTAATTAGAGATTAGAAGAAAATTAAGGATGGCAGT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 TTGCTTGAAGAAGAGATAATGGATTAAACTAATTAGAGAATTAGAAGAAAATTAAGGATGGAGA 270

Qy    349 ATCACACATAATAAAAGTTTCTACCTTTCTTAAAAGTTGAGTTGTTTTTTAAATTG 408
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 GTCACACATAATAAAGTTTCTACCTTTCTTAAAAGTGTAGTTATGTTTTTTAAATTG 330

Qy    409 GCTTTTAGTTAGTCCTTTATTACACTTTAAATAAATAAGTGTGGGCACTTCAAATCCTTTGTGCAAT 468
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 GCTTTTAGTTAGTCCTTTTATTACACTTTAAATAAATAAGTGTGGGCACTTCAAACCCTTTGTGCAAT 390

Qy    469 CTTGCACTAAGTTTATTTTTGTACTTTTAAATGAAAAATGAACTTCTATGTCTTTGGCTTAA 528
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 CTTGCACTAAGTTTATTCGTGTACTTTTAATGAAAAATCACCCTTCTATGGTTTTTGTTTAA 450
```

```
Qy    529 AAAAAA 534
          ||||||
Db    451 AAAAAA 456
```

```
RESULT 4'
ID   AAQ10262 standard; cDNA; 564 BP.
XX
AC   AAQ10262;
XX
DT   25-MAR-2003  (updated)
DT   04-APR-1991  (first entry)
XX
DE   Ovary tissue transcriptional factor DNA clone p2130.
XX
KW   Ovary tissue transcriptional factor; DNA construct; probe;
KW   clone p2130; ds.
XX
OS   Lycopersicon esculentum UC82B.
XX
PN   EP409629-A.
XX
PD   23-JAN-1991.
XX
PF   19-JUL-1990;  90EP-0307926.
XX
PR   19-JUL-1989;  89US-0382518.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   Martineau B, Houck CM;
XX
DR   WPI; 1991-024191/04.
DR   P-PSDB; AAR10310.
XX
PT   New ovary tissue transcriptional factors - modify transcription
PT   in tomato plant ovaries for use as mol. probes
XX
PS   Disclosure; Fig 1; 21pp; English.
XX
CC   The tomato-derived transcriptional initiation region which regulates
CC   the expression of the sequence corresp. to the p2130 clone is
CC   considered ovary-specific. Sequences hybridisable to the p2130
CC   clone, e.g. probe p27, show abundant mRNA, esp. at the early stages
CC   of anthesis. The message is expressed in ovary integument and ovary
CC   outer pericarp tissue and is not expressed, or at least not readily
CC   detectable, in other tissues or at any other stage of fruit
CC   development. The native function of the p2130 amino acid sequence
CC   encoded by the structural gene comprising p2130 is unknown.
CC   See also AaQ10263-64.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;
```

```
FH   Key             Location/Qualifiers
FT   misc_feature    447..564
FT                   /*tag= a
FT                   /label= p27_probe
```

Query Match   35.1%; Score 189.8; DB 12; Length 564;
Best Local Similarity 68.9%; Pred. No. 6.5e-37;
Matches 326; Conservative   0; Mismatches 127; Indels 20; Gaps 4;

```
Qy      1 ATGGGTCGCCCCTTGTGCTTCATGGCCATTGCTATCTTGGCCAAGGATGCTCTTTGTTGCC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 ATGGGCTCGTTCCATTTTCTTCATGGCATTTTTTGGTCTTTGGCATTGCTGAAGTCTTTGTTTACC 99

Qy     61 TATGAGGTGCAAGCTGTAGAA----TGCAAAACAGAAAGCAACACACATTTCCTGGAATATGC 117
          |||||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db    100 TATGAGGTAGAAGCTCAGCAAATTTGCAAAGCACCAAGCCAAACTTTCCCAGGAATTATGT 159

Qy    118 ATTACCAAAACCACCATGCGAGAAAAGCTTGTATCAGTGAGAAAATTTACTGATGGGGTCATTGT 177
          |||||||||||||||||||||||||||||||||||||||||||
Db                                                            (continues)
```

RESULT 5

AAQ34940

AAQ34940 standard; DNA; 564 BP.

ID   AAQ34940 standard; DNA; 564 BP.
XX
AC   AAQ34940;
XX
DT   25-MAR-2003 (updated)
DT   19-MAY-1993 (first entry)
XX
DE   pz130 coding sequence.
XX
KW   cDNA; clone pz130; anthesis; tomato; ovary; integumen; outer pericarp;
KW   fruit; developement; transcription; initiation; region; modulation;
KW   ovary-specific; endogenous; fruit product; exogenous; phenotype; ds.
OS   Lycopersicon esculentum.
XX
FH   Key   Location/Qualifiers
FT   misc_RNA   447..564
FT   /*tag=   a
FT   /note= "Corresponds to p27 cDNA".
XX

Qy   457 TATTGGCCTATAGTAGTAGCCATTTGACACATTAAATAAGTTTGTGACACAATCATT 509

Db   457 TATTGGCCTATAGTAGTAGCCATTTGACACATTAAATAAGTTTGTGACACAATCATT 509

Qy   404 AATTGGCTTTAGTAGTAGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456
Db   404 AATTGGCTTTAGTAGTAGCCTTTTATTACACTTTAAATAAGTGTGGCACTTCAAT 456

Qy   397 AAAACAAAAATTAATAAAGTGTGCCTTTTCTTATTAGGGTTAGTTGGC---TAAATGTTGTGTTTT 403
Db   397 AAAACAAAAATTAATAAAGTGTGCCTTTTCTTATTAGGGTTAGTTGGC---TAAATGTTGTGTTTT 456

Qy   347 GTATCACACATAATAAAGTTTCTACCTTTTCTTAAAAAGTGTAGC--- TAAATGTTGTGTTTT 403
Db   347 GTATCACACATAATAAAGTTTCTACCTTTTCTTAAAAAGTGTAGC--- TAAATGTTGTGTTTT 456

Qy   337 GAAGAGATTATGCATGGAGTAATAATTAGAGAATTTTGCAGGTGTTAAA 396
Db   337 GAAGAGATTATGCATGGAGTAATAATTAGAGAATTTTGCAGGTGTTAAA 403

Qy   298 GAAGAGATAATGGATAAACTAATTAGAGAAGAAAATTAAGGAT-------GCA 346
Db   298 GAAGAGATAATGGATAAACTAATTAGAGAAGAAAATTAAGGAT-------GCA 396

Qy   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336
Db   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336

Qy   238 AAAACAGGAGCTGAAATTTGGCTCGAGGAAGCAAAAACTTTGGCTGCACGCTTTGCTTGAA 297
Db   238 AAAACAGGAGCTGAAATTTGGCTCGAGGAAGCAAAAACTTTGGCTGCACGCTTTGCTTGAA 297

Qy   220 AGCAAACTCCAAAGGAAGGTGTCTATGCACTAAGCCATGCATGTGTGATT---TGACAAAATCTCA 276
Db   220 AGCAAACTCCAAAGGAAGGTGTCTATGCACTAAGCCATGCATGTGTGATT---TGACAAAATCTCA 276

Qy   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGCATGTGTGTTGATGAGAAGAAGATGACT 237
Db   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGCATGTGTGTTGATGAGAAGAAGATGACT 237

Qy   160 TTTATGGACTCATCATTGTAGAAAATATTGTATCAAAGAGAAAATTTACTGGTGGACATTGT 219
Db   160 TTTATGGACTCATCATTGTAGAAAATATTGTATCAAAGAGAAAATTTACTGGTGGACATTGT 219

SQ   Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;

Query Match   35.1%; Score 189.8; DB 14; Length 564;
Best Local Similarity   68.9%; Pred. No. 6.5e-37;
Matches 326; Conservative   0; Mismatches 127; Indels 20; Gaps 4;

CC   associated with this gene is therefore considered to be ovary-
CC   specific. The actual function of the pz130 polypeptide is unknown.
CC   The transcription initiation region can be used for modulation of
CC   endogenous fruit products, for production of exogenous products and
CC   for modification of the phenotype of fruit and fruit products.
CC   (Updated on 25-MAR-2003 to correct PF field.)

RESULT 6

AAT31823

AAT31823 standard; cDNA; 564 BP.

ID   AAT31823 standard; cDNA; 564 BP.
XX
AC   AAT31823;
XX
DT   25-MAR-2003 (updated)
DT   14-SEP-1996 (first entry)
XX
DE   Tomato pz130 cDNA clone.
XX
KW   Ovary; ovule; fruit; tomato; cotton; melanin; vector;
KW   transgenic plant; ss.
OS   Lycopersicon esculentum cv. UC82B.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..357
FT   /*tag=   a
XX

Qy   457 TATTGGCCTATAGTAGTAGCCATTTGACACATTAAATAAGTTTGTGACACAATCATT 509
Db   457 TATTGGCCTATAGTAGTAGCCATTTGACACATTAAATAAGTTTGTGACACAATCATT 509

Qy   397 AAAACAAAAATTAATAAAGTGTGCCTTTTCTTATTAGGGTTAGTTGGCTTGTGAGTGTTGTGTTTAG 456
Db   397 AAAACAAAAATTAATAAAGTGTGCCTTTTCTTATTAGGGTTAGTTGGCTTGTGAGTGTTGTGTTTAG 456

Qy   347 GTATCACACATAATAAAGTTTCTACCTTTTCTTAAAAAGTGTAGC---TAAATGTTGTGTTTT 403
Db   347 GTATCACACATAATAAAGTTTCTACCTTTTCTTAAAAAGTGTAGC---TAAATGTTGTGTTTT 403

Qy   298 GAAGAGATAATGGATAAACTAATTAGAGAATTAGAGATTAGAGAAAATTAAGGAT 346
Db   298 GAAGAGATAATGGATAAACTAATTAGAGAATTAGAGATTAGAGAAAATTAAGGAT 346

Qy   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336
Db   277 AGTGAAGTTAAAGCAACTTTGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAA 336

Qy   238 AAAACAGGAGCTGAAATTTGGCTCGAGGAAGCAAAAACTTTGGCTGCACGCTTTGCTTGAA 297
Db   238 AAAACAGGAGCTGAAATTTGGCTCGAGGAAGCAAAAACTTTGGCTGCACGCTTTGCTTGAA 297

Qy   220 AGCAAACTCCAAAGGAAGGTGTCTATGCACTAAGCCATGCATGTGTGATT---TGACAAAATCTCA 276
Db   220 AGCAAACTCCAAAGGAAGGTGTCTATGCACTAAGCCATGCATGTGTGATT---TGACAAAATCTCA 276

Qy   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGCATGTGTGTTGATGAGAAGAAGATGACT 237
Db   178 AGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGCATGTGTGTTGATGAGAAGAAGATGACT 237

Qy   160 TTTATGGACTCATCATTGTAGAAAATATTGTATCAAAGAGAAAATTTACTGGTGGACATTGT 219
Db   160 TTTATGGACTCATCATTGTAGAAAATATTGTATCAAAGAGAAAATTTACTGGTGGACATTGT 219

Qy   100 TATGGAGGTAGAAGCTCAGACAAATTTGCAAAGACCACCAAGCCAAAACTTTCCCAGGATTATGT 159
Db   100 TATGGAGGTAGAAGCTCAGACAAATTTGCAAAGACCACCAAGCCAAAACTTTCCCAGGATTATGT 159

Qy   61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGGAAGCAACACAATTTCCTGGAAATGC 117
Db   61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGGAAGCAACACAATTTCCTGGAAATGC 117

Qy   40 ATGGCTCGTTCCAATTTCTTCAATGGCATTTGGCAAATGAGTTGTGCTTTTGTTACC 99
Db   40 ATGGCTCGTTCCAATTTCTTCAATGGCATTTGGCAAATGAGTTGTGCTTTTGTTACC 99

Qy   1 ATCGGCTCGGCTCCGTTCGTCCTTTCTTCATTGTTGGCTGCAGCTTTGCTTGAA 60
Db   1 ATCGGCTCGGCTCCGTTCGTCCTTTCTTCAATTGTTGGCTGCAGCTTTGCTTGAA 60

Qy   1 ATCGGCTCGGCTCCGTTCGTCCTTTCTTCAATTGTTGGCTGCAGCTTTGCTTGAA 60
Db   1 ATCGGCTCGGCTCCGTTCGTCCTTTCTTCAATTGTTGGCTGCAGCTTTGCTTGAA 60

PN   US5530185-A.
XX
PD   29-DEC-1992;
PF   29-JUN-1996.
PR   29-DEC-1992;   92US-0998158.
PR   92US-0998158.

PN   US5175095-A.
XX
PD   29-DEC-1992.
XX
PF   17-JUL-1990;   90US-0554195.
PR   19-JUL-1989;   89US-0382518.
PR   17-JUL-1990;   90US-0554195.
XX
PA   (CALJ) CALGENE INC.
XX
PI   Houck CM, Martineau BM;
XX
DR   WPI; 1993-026940/03.
DR   P-PSDB; AAR30779.
XX
PT   DNA constructs contg. tomato p2130 transcriptional initiation
PT   region - useful for modulation of endogenous fruit prods. and for
PT   prodn. of exogenous prods.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence given shows the DNA sequence of cDNA clone p2130. This
CC   sequence is expressed during the early stages of anthesis in tomatos.
CC   The message is expressed in ovary integumen and ovary outer pericarp
CC   region. of exogenous prods.
CC   tissue. It is not readily detectable in other tissues or at other
CC   stages of fruit developement. The transcription initiation region

```
PR   19-JUL-1989;   89US-0382518.
PR   17-JUL-1990;   90US-0554195.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   Martineau BM, Reilley AA,  Stalker DM;
XX
DR   WPI; 1996-308822/31.
DR   P-PSDB; AAR97559.
XX
SQ   Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;
```

Query Match   35.1%; Score 189.8; DB 17; Length 564;
Best Local Similarity   68.9%; Pred. No. 6.5e-37;
Matches 326; Conservative   0; Mismatches 127; Indels 20; Gaps 4;

```
Qy      1 ATGGCTCGCTCCTTGTTGCTTCANGGCANTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC     60
          ||||||||||||||||||||||||| ||||| || |||||||||||||||||||||||||||
Db     40 ATGGCCTCGTTCCATTTTCTTCANGGCCANTTTTNGGTCTTGGCAATGAGCTCTTTGTTACC     99

Qy     61 TATGAGGTGCAAGCTAGAGAA---TGCAAAAACAGAAAAGCAACACANTTTCCTGGAATATGC    117
          ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||
Db    100 TATGAGGTAGAAGCTCAGCAAAATTTGCAAAAGCACCAAACTTTCCCAGGATTATGT    159

Qy    118 ATTACCAAACCACCCATGCAGAAAAAGCTTGTATCAGTGAGAAAATTTACTGAGGTCATTGT    177
          |||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    160 TTTATGGACTCATCATGTAGAAAAATTTATCAAAGAGAAAATTTACTGGTGGACATTGT    219

Qy    178 AGCAAAAATCCTCAGAAGGTGCCTAGTACTAAGCCCATGTGTGTTGATGAGAAGATGACT    237
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 AGCAAAACTCCAAAGGAAGTGTCTATCCACTAAGCCATGTGTATT---TGACAAAAATCTCA    276

Qy    238 AAAACAGGAGCTGAAATTTGGCTGAGGAAGCAAAAAACTTTGGCTGCAGCTTTGCTTGAA    297
          ||         ||||||||||||||||||||||||||||||||||||
Db    277 AGTGAAGTTAAAGCAACTTCGGGTGAGGAAGCAAAAACTCTAAGTGAAGTTGGTGCTTGAA    336

Qy    298 GAAGAGAATAATGGATAACTAATTAGAGATTAGAAGGANATTAAGGA?----------GCA    346
          ||||||||||||||||||||
Db    337 GAAGAGAATTATGCATGGAGAGTAAATAATTAAGTGAGGTTAAAATAAGGATTTGAGTGTGTCAAAA    396

Qy    347 GTATCACACATAATAATAAAGTTTCTACTT----------TAANGTTGTTGTTTT    403
          ||||||||||||||||
Db    397 AAAACAAAAATTAATAAAGTGTTGCCCTTTTCTTATTATTAAAAGTGTAGC--    456

Qy    404 AATTGGCTTTAGTTAGTAGCCTTTTATTTACACTTTAAATATAAGTGGTGGCACTTCAAT    456
          |||||||||||||||||||||||||||||||
Db    457 TATTGGCCTAGTAGTAGCCCATTTAAATAAGTTTGTGTGACACATCATT    509
```

Left column

```
Qy    298  GAAGAGATAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGAT----------------GCA 346
                 ||||||  ||||  ||  ||  |||  ||  |||   |||||  ||
Db    337  GAAGAGATTATGGAGAGTGGAGTAATAATTAAGGAGGTTAAATAAGGAGATTTGAGTGTCAAA 396

Qy    347  GTATCACACATAAATAAAGTTTCTACCTTTCTTAAAAGTGTAGC---TAAGTGTGTGTTTT 403
                 ||||  |||  ||  |||||||
Db    397  AAAACAAAATTAATTAAAGTGTTGCCTTTTCTTATTAGGGTAGCAGTTGATGTGTGTTAG 456

Qy    404  AATTGGGCTTTTAGTTAGCCTTTTATTACACTTTAAATAATAAGTGTGGCCACTTCAAT 456
                 |||||||  |  ||||||||||  |  | |  | |||  |||  |  |
Db    457  TATTGGGCCTATAGTTAGCCCATTTGACACATTAAATAATAAGTTTGTGACACATCATT 509
```

RESULT 8

AB221970

ID   AB221970 standard; cDNA; 564 BP.
XX
AC   AB221970;
XX
DT   28-MAR-2003 (first entry)
XX
DE   Plasmid pz130 cDNA sequence p27.
XX
KW   Transcriptional factor; ovary tissue; gene; ss.
XX
OS   Unspecified.
XX
FH   Key   Location/Qualifiers
FH
CDS   1..564
     /partial
     /*tag= a
     /product= "pz130 protein sequence"
     /trans_except= (pos:355..357,aa:Xaa)
     /trans_except= (pos:358..360,aa:Xaa)
     /trans_except= (pos:376..378,aa:Xaa)
     /trans_except= (pos:385..387,aa:Xaa)
     /trans_except= (pos:436..438,aa:Xaa)
     /trans_except= (pos:454..456,aa:Xaa)
     /trans_except= (pos:478..480,aa:Xaa)
     /trans_except= (pos:490..492,aa:Xaa)
     /trans_except= (pos:541..543,aa:Xaa)
     /note= "all Xaa's are encoded by stop codons"
XX
CN1189856-A.
XX
PN   CN1189856-A.
XX
PD   05-AUG-1998.
XX
PF   07-JUN-1996; 96CN-0195170.
XX
PR   07-JUN-1995; 95US-0480087.
XX
PA   (CALJ) CALGENE INC.
XX
DR   WPI; 2002-733392/80.
DR   P-PSDB; ABP56255.
XX
PT   Use of transcriptional factors of ovary tissue -
PS   Example 3; Fig 1A-B; 56pp; Chinese.
XX
CC   The present invention describes the use of transcriptional factors of
CC   ovary tissue. The present sequence represents a nucleotide sequence
CC   which is used in an example from the present invention.
XX
SQ   Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;

Query Match   35.1%;   Score 189.8; DB 24; Length 564;
Best Local Similarity 68.9%; Pred. No. 6.5e-37;
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

```
Qy   1  ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column

```
Db    40  ATGGCTCGCTTCCATTTTCTTCATGGCATTTTGGTCTTGGCAAATGATGCTCTTTGTTACC 99

Qy    61  TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAGAAGCACACATTTCCTGGAAATGC 117
              ||||||||||||||||||||      ||  ||  | |||||||  ||  |  ||||| ||
Db   100  TATGAGGTTAGAAGCTCAGCAAATTGCAAAAACTCAGCAAAGCCAAACTTTCCCACGAATTATGT 159

Qy   118  ATTACCAAACCACCACATGCAGAAAAAGCTGTATCAGTGAGAAATTACTGATGGTCATTGT 177
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db   160  TTTATGGACTCATCATGTAGAAAAATATTGTATCAAAATGAGAAATTTACTGGTGGACATTGT 219

Qy   178  AGCAAAATCCTAGAAAGGTGCCTAGTTACTAAGCCATGTGTGTTGATGAGAGAGATGACT 237
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220  AGCAAAACTCCAAAGGAAGTGTCTATCGACTAAGCCATGTGTATT---TGACAAAATCTCA 276

Qy   238  AAAACAGGAGCTGAAAATTTTGGCTGAGGAGGAAGCAAAAAACTTTGGCTGCGAGCTTTGCTTGAA 297
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277  AGTGAGGTTAAAGCAACTTTGGGGTGAGGAAGCAAAAAACTCTAAGTGAAAGTGTGTGCTTGAA 336

Qy   298  GAAGAGATAATGGATAACTAATTAGAGATTAGAAGAAAATTAAGGAT------------GCA 346
                 ||||||  ||||  ||  ||  |||  ||  |||   |||||  ||
Db   337  GAAGAGATTATGGAGAGTGGAGTAATAATTAAGGAGGTTAAATAAGGAGATTTGAGTGTCAAAA 396

Qy   347  GTATCACACATAAATAAAGTTTCTACCTTTCTTAAAAGTGTAGC---TAAGTGTGTGTTTT 403
                 ||||  |||  ||  |||||||
Db   397  AAAACAAAATTAATTAAAGTGTTGCCTTTTCTTATTAGGGTAGCAGTTGATGTGTGTTAG 456

Qy   404  AATTGGGCTTTTAGTTAGCCTTTTATTACACTTTAAATAATAAGTGTGGCCACTTCAAT 456
                 |||||||  |  ||||||||||  |  | |  | |||  |||  |  |
Db   457  TATTGGGCCTATAGTTAGCCCATTTGACACATTAAATAATAAGTTTGTGACACATCATT 509
```

RESULT 9

AAT31824

ID   AAT31824 standard; cDNA; 3528 BP.
XX
AC   AAT31824;
XX
DT   25-MAR-2003 (updated)
DT   14-SEP-1996 (first entry)
XX
DE   Tomato pz130 genomic clone Calgene Lambda 140.
XX
KW   Ovary; ovule; fruit; tomato; cotton; melanin; vector;
KW   transgenic plant; ss.
XX
OS   Lycopersicon esculentum cv. UC82B.
XX
FH   Key   Location/Qualifiers
FH
misc_difference 586
     /*tag= a
     /note= "base n at position 586 is unidentified"
misc_difference 590
     /*tag= b
     /note= "base n at position 590 is not identified"
misc_difference 591
     /*tag= c
     /note= "base n at position 591 is not identified"
misc_difference 663
     /*tag= d
     /note= "base n at position 663 is not identified"
misc_difference 667
     /*tag= e
     /note= "base n at position 667 is not identified"
misc_difference 769
     /*tag= f
     /note= "base n at position 769 is not identified"
misc_difference 2483
     /*tag= g
     /note= "base 2483 is given as s in the
            specification"
misc_difference 2485
     /*tag= h
     /note= "base 2485 is given as s in the

FT   specification"
FT   exon   2567..2701
FT   /*tag= i
FT   intron   2702..2921
FT   /*tag= j
FT   exon   2922..3528
FT   /*tag= k
XX
PN   US5530185-A.
XX
PD   25-JUN-1996.
XX
PF   29-DEC-1992; 92US-0998158.
XX
PR   29-DEC-1992; 92US-0998158.
PR   19-JUL-1989; 89US-0382518.
PR   17-JUL-1990; 90US-0554195.
XX
PA   (CALJ) CALGENE INC.
XX
PI   Martineau BM, Reilley AA, Stalker DM;
XX
DR   WPI; 1996-308822/31.
DR   P-PSDB; AAR97559.
XX
PS   Example 5; Fig 2A-C; 25pp; English.

CC   The tomato pZ130 genomic clone (AAT31824) was isolated from a genomic
CC   library using a pZ130 cDNA clone (AAT31823) as probe. The pZ130
CC   transcriptional initiation region is considered to be ovary-specific.
CC   It can be utilised in DNA constructs for the expression of
CC   heterologous genes, partic. in early fruit development, and esp. for
CC   expression of a melanin synthesis gene in transgenic cotton. The
CC   native function of the pZ130 gene product (AAR97559) is unknown.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX

DNA construct for expressing melanin synthesis gene in plant ovule
cells - contains promoter from the tomato pZ130 gene, also binary
vector and transgenic plants, esp. cotton, contg. construct

SQ   Sequence 3528 BP; 1191 A; 529 C; 1282 T; 8 other;

Query Match   27.2%; Score 147.4; DB 17; Length 3528;
Best Local Similarity   67.1%;
Matches 261; Conservative   0; Mismatches 111; Indels 17; Gaps 3;
                         Pred. No. 2.3e-26;

Qy   82 TGCAAAACAGAAAGCAACACACATTTCCTGGAATATGCATTACCAAAACCACCATGCAGAAAA 141
          |||||||||| ||||| ||| |||||| |||||| |||||| |||| ||||| ||||||
Db   2942 TGCAAACGACCAAGCCAAACTTTCCCAGGATTATGTTTATTGTTTAATGGACTCATCATGTAGAAAA 3001

Qy   142 GCTTGTATCAGTGAGAAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
          |||||||||| |||||||||||||| ||| |||||||||||||||||| ||| ||||
Db   3002 TATTGTATCAAAGAGAAAATTTACTGGTCGGACATTGTAGCAAAATCCTCAAAGGAAGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTGTTTGATGAGAGAAGATGACTAAAAACAGGAGCTGAAATTTTGGGCT 261
          || |||||||||||||| || ||| ||||| |||| ||||| ||||||| ||||
Db   3062 TGCACTAAGCCATGTGTATT--TGACAAAATCCAAGTGAGAGTTAAAGCAACACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAAACTTTGGCTCGCAGCTTTGCTTGACGAAGAGAGATAATGGATAACTAATTA 321
          |||||||||||||||| |||||||||| ||||||||||| |||||||| ||||| || ||||
Db   3119 GAGGAAGCAAAAAACTCTAAGTGAAGTTGCTGCTTGAGAGAAGAGATTATGATGGAGTAATAA 3178

Qy   322 GAGAGTAGAAGAAAATTAAGGAT------------------GCAGTATCACACATAATAAAAGTTTCTA 370
          ||| ||| ||||||||||||| |||||||||| |||||| ||| ||||
Db   3179 TTAAGTGAGGTTAAATAAGGAT   TTTGAGTGTCAAAAAAAAACAAAATTAATAAAGTCGTTGC 3238

Qy   371 CCTTTCTTATTTAAAAGTGTAGC---TAATCGTTGTGTTTTAATTGGCTTTTAGTTAGTAGCCTTTTAT 427
          | ||||||||| |||| |||||| || ||| |||||| ||||||| ||||||||||| ||||
Db   3239 CTTTTCTTATTAGGGTAGCTTCGTTCGTTAGTATTGGCCCTAAGTAGCCCATTTGA 3298

Qy   428 TACACTTTAAATAAGTTTGGACATGGCCACTCCAAT 456
          | |||| |||||||||||||||| ||| |||||
Db   3299 CACATTAAATAAGTTTGGACACATCAAT 3327

RESULT 10
AAT48816
ID   AAT48816 standard; cDNA; 3528 BP.
XX
AC   AAT48816;
XX
DT   14-MAR-1997 (first entry)
XX
DE   Calgene Lambda 140 genomic clone.
XX
KW   Ovary; fruit; colour; pigmentation; cotton; tomato; probe;
KW   promoter; pZ130; pZ7; transgenic plant; ss.
XX
OS   Lycopersicon esculentum cv. UC82B.
XX
PN   WO9640951-A2.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996; 96WO-US09911.
XX
PR   07-JUN-1995; 95US-0487087.
XX
PA   (CALJ) CALGENE INC.
XX
PI   McBride K, Stalker DM;
XX
DR   WPI; 1997-052341/05.
XX
PS   Example 7; Fig 2; 75pp; English.

DNA construct capable of directing ovary-tissue transcription in
plants - useful for modifying colour phenotype, in e.g. cotton

CC   Calgene Lambda 140 genomic clone (AAT48816) overlaps with the
CC   tomato ovary-specific pZ130 cDNA clone (see also AAT48813) and
CC   includes regions 5' and 3' to pZ130. It was isolated from a
CC   tomato cv. UC82B genomic library using pZ130 as probe. The
CC   genomic clone can be used to construct pZ130 promoter cassettes
CC   useful for directing ovary tissue transcription in plants, e.g. to
CC   modify colour phenotype in cotton.
XX
SQ   Sequence 3528 BP; 1191 A; 529 C; 1282 T; 8 other;
XX
FH   Key   Location/Qualifiers
FH
FT   exon   2567..2701
FT   /*tag= a
FT   intron   2702..2921
FT   /*tag= b
FT   exon   2922..3528
FT   /*tag= c
FT   misc_RNA   2599..2701
FT   /*tag= d
FT   /note= "5' end of pZ130 gene"
FT   misc_RNA   2922..3382
FT   /*tag= e
FT   /note= "3' end of pZ130 gene"

Query Match   27.2%; Score 147.4; DB 18; Length 3528;
Best Local Similarity   67.1%;
Matches 261; Conservative   0; Mismatches 111; Indels 17; Gaps 3;
                         Pred. No. 2.3e-26;

Qy   82 TGCAAAACAGAAAGCAACACACATTTCCTGGAATATGCATTACCAAAACCACCATGCAGAAAA 141
          |||||||||| ||||| ||| |||||| |||||| |||||| |||| ||||| ||||||
Db   2942 TGCAAACGACCAAGCCAAACTTTCCCAGGATTATGTTTATTGTTTAATGGACTCATCATGTAGAAAA 3001

Qy   142 GCTTGTATCAGTGAGAAAATTTACTGATGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
          |||||||||| |||||||||||||| ||| |||||||||||||||||| ||| ||||
Db   3002 TATTGTATCAAAGAGAAAATTTACTGGTCGGACATTGTAGCAAAATCCTCAAAGGAAGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTGTTTGATGAGAGAAGATGACTAAAAACAGGAGCTGAAATTTTGGGCT 261
          || |||||||||||||| || ||| ||||| |||| ||||| ||||||| ||||
Db   3062 TGCACTAAGCCATGTGTATT--TGACAAAATCCAAGTGAGAGTTAAAGCAACACTTTGGGT 3118

Db 3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy 262   GAGGAAGCAAAAACTTTGGCTGCACTTTGCTTGAGAGAGAGAAGTTGCTTGAGAGAGAGA 321

Db 3119 GAGGAAGCAAAAACTTCAAGTGAAGTTGTGCCTTGAGAAGAGAGATAATGATACTAATTA 3178
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy 322   GAGATTAGAAGAAATTAAGGAT------------GCAGTATCACACATAATAAATAAGTTCTA 370

Db 3179 TTAAGTGAGGTAAATAAGGATTTTGAGTGTCAAAAAAAACAAAATAATAATAAAGTGTTGC 3238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 371   CCTTTCTTAAAAGTCTAGC---TAATGTTCTGTTGTTTTAATTGGCTTTTTAT 427

Db 3239 CTTTTCTTATTAGGGTAGCTTGTGTGAGTGTGTTAGTATTGGCCTATAGTAGCCATTTGA 3298
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 428   TACACTTTAAATAAGTGTGGCACTTCAAT 456

Db 3299 CACATTAAAATAAGTTTGTGACACATCATT 3327

RESULT 11
ID   ABZ21971 standard; cDNA; 3528 BP.
XX
AC   ABZ21971;
XX
DT   28-MAR-2003 (first entry)
XX
DE   Plasmid pZ130 cDNA sequence calgene lambda 140.
XX
KW   Transcriptional factor; ovary tissue; gene; ss.
XX
OS   Unspecified.
XX
FH
PN   CN118956-A.
XX
PD   05-AUG-1998.
XX
PF   07-JUN-1996;   96CN-0195170.
PR   07-JUN-1995;   95US-0480087.
XX
PA   (CALJ) CALGENE INC.
XX
DR   WPI; 2002-733392/80.
XX
PT   Use of transcriptional factors of ovary tissue -
XX
PS   Example 5; Fig 2A-C; 56pp; Chinese.
XX
CC   The present invention describes the use of transcriptional factors of
CC   ovary tissue. The present sequence represents a nucleotide sequence
CC   which is used in an example from the present invention.
XX
SQ   Sequence 3528 BP; 1191 A; 529 C; 1282 T; 8 other;

Query Match   27.2%; Score 147.4; DB 24; Length 3528;
Best Local Similarity   67.1%; Pred. No. 2.3e-26;
Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

Qy 82   TGCAAAAACAGAAAGCAACACATTTCTGGAATATGCATTACCAAACCACCATGACAGAAA 141
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2942 TGCAAAGCACCAAACTTTACTGAAAGATTTTCCAGGATTATGTTTTATGGACTCATCACATTTTTATGGACTCATCACATTTCCAGGATTATGTTTTATGGACTCATCACA 3001

Qy 142   GCTTGTTGTATCAGTGAGAAATTTACTGATGAGTGGTCAATGCTGAGCAAAAATCCTCAGAAGTGCCTA 201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3002 TATTGTTATCAAAGAGAAATTTACTCGTGACATTGTTAGCAAAACTCCAAAGAAGTGTCTA 3061

Qy 202   TGTACTTAAGCCATGTGTTGTTTGATGAGGAAGATGACTAAAAACAGGAGCTGAAATTTTGGCT 261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy 262   GAGGAAGCAAAAACTTTGGCTGCACTTTGCTTGAGAGAGAGAAGTTGCTTGAGAGAGAGAGAAGATTATGATGGAGTAATAA 321

RESULT 12
ID   AAQ10263 standard; DNA; 4383 BP.
XX
AC   AAQ10263;
XX
DT   25-MAR-2003 (updated)
DT   04-APR-1991 (first entry)
XX
DE   pZ130 contg. Calgene lambda 140 genomic clone.
XX
KW   Ovary tissue transcriptional factor; DNA construct; probe;
KW   clone pZ130; ss.
XX
OS   Lycopersicon esculentum UC82B.
XX
FH   Key   Location/Qualifiers
FT   intron   2702..2921
FT   /*tag= a
FT   2599..3382
FT   /*tag= b
FT   misc_RNA   /label= pZ130
FT   /note= "overlap"
FT   misc_feature   2567..2567
FT   /*tag= c
FT   /label= transcription_start
XX
PN   EP409629-A.
XX
PD   23-JAN-1991.
XX
PF   19-JUL-1990;   90EP-0307926.
PR   19-JUL-1989;   89US-0382518.
XX
PA   (CALJ) CALGENE INC.
XX
PI   Martineau B, Houck CM;
XX
DR   WPI; 1991-024191/04.
XX
PT   New ovary tissue transcriptional factors - modify transcription
PT   in tomato plant ovaries for use as mol. probes
XX
PS   Disclosure; Fig 2; 21pp; English.
XX
CC   To prepare pCGN2901, Calgene Lambda 140 was digested with SalI
CC   and the resulting fragment which contains the pZ7-hybridising
CC   region was inserted into pCGN2015, at the unique SalI site.
CC   pCGN2902 contains the other SalI fragment (non-pZ7-hybridising)
CC   of the pZ130 genome derived from SalI digestion of
CC   Calgene lambda 140 also put into a pCGN2015 construct.
CC   Plasmid DNA isolated from pCGN2901 was used in the construction
CC   of a pZ130 Promoter Cassette. The promoter is ovary-specific
CC   and wound-inducible in leaf cells.
CC   See also AAQ10262-64.
CC   (Updated on 25-MAR-2003 to correct PA field.)

Db 3119 GAGGAAGCAAAAACTTCAAGTGAAGTTGTGCCTTGAGAAGAGAGATAATGATACTAATTA 3178
Qy 322   GAGATTAGAAGAAATTAAGGAT------------GCAGTATCACACATAATAAATAAGTTCTA 370
Db 3179 TTAAGTGAGGTAAATAAGGATTTTGAGTGTCAAAAAAAACAAAATAATAATAAAGTGTTGC 3238
Qy 371   CCTTTCTTAAAAGTCTAGC---TAATGTTCTGTTGTTTTAATTGGCTTTTTAT 427
Db 3239 CTTTTCTTATTAGGGTAGCTTGTGTGAGTGTGTTAGTATTGGCCTATAGTAGCCATTTGA 3298
Qy 428   TACACTTTAAATAAGTGTGGCACTTCAAT 456
Db 3299 CACATTAAAATAAGTTTGTGACACATCATT 3327

XX
SQ Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;

Query Match   27.2%; Score 147.4; DB 12; Length 4383;
Best Local Similarity   67.1%; Pred. No. 2.5e-26;
Matches 261; Conservative   0; Mismatches 111; Indels 17; Gaps   3;

Qy   82 TGCAAAACAGAAGAGCAACACATTTCCTGGAATATGCATTACCAAACCACCATGCAGAAAA 141
         |||||| || || |||| ||||| |||| ||||| ||| || ||| |||||| ||||
Db 2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTTATGGACTCATCATGTAGAAAA 3001

Qy   142 GCTTGTATCAGTGAGAAAATTTACTGATGCTCATTGTTAGCAAAATCCTCAGAAGGTGCCTA 201
         || ||||||| ||||||||||||| | | |||||||||||||| ||| || |||||| ||||
Db 3002 TATTGTATCAAAGAGAAATTTACTCGGTGGACATTGTTAGCAAAACTCCAAAGAAGGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTGTTTGATGAGGAAGATGACTAAAAACAGGAGCTGAAAATTTTGGCT 261
         || ||||||||| |||| || ||| |||| ||| | ||| || ||| |||| |
Db 3062 TGCACTAAGCCAATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAAGAGAGATAATGGATTAACTAATTA 321
         ||||||||||||| | || | |||| ||||||||||||||||||| ||| |||| | ||
Db 3119 GAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAAGAAGAGAGATTATGTTTTATGGGGAGTAATAA 3178

Qy   322 GAGATTAGAAGAAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTTTCTA 370
         ||| ||| ||||| |||||| ||||||| |||||||||||||||| | |
Db 3179 TTAAGTGAGGTTAAATAAGGAGATTTTGAGTGTCAAAAAAAACAAAAATTAAATAAAAGTGTTGC 3238

Qy   371 CCTTTCTTAAAAAGTGTAGC---TAAGTGTTGTTTTTAAATTGGCTTTTAGTTAGTTAGCCTTTTAT 427
         | ||||||| |||| ||| | || ||||| ||| ||| ||||| ||||||| ||
Db 3239 CTTTTCTTATTAGGGTGAGCTTGTGTCGTTGTTAGTTAGTTAGTAGCCATTGA 3298

Qy   428 TACACTTTAAATAAGTTTGTGTGGCACTTCAAT 456
         | |||| |||||||||||||| || | | ||
Db 3299 CACATTAAATAAGTTTGTGACACATCATT 3327

RESULT 13
AAQ10319

ID AAQ10319 standard; cDNA; 4383 BP.
XX
AC AAQ10319;
XX
DT 25-MAR-2003 (updated)
DT 04-APR-1991 (first entry)
XX
DE Calgene lambda 140 genomic clone.
XX
KW pz130; cytokinin; ds.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_RNA   2567..3382
FT   /*tag= a
FT misc_RNA   2599..3382
FT   /*tag= b
FT   /note= "Sequence homologous to pz130 transcript"
FT misc_RNA   2702..2921
FT   /note= "Sequence homologous to pz130 CDS"
FT   /*tag= c
FT   /note= "Sequence homologous to pz130 intron"
XX
PN EP409628-A.
XX
PD 23-JAN-1991.
XX
PF 19-JUL-1990; 90EP-0307925.
XX
PR 19-JUL-1989; 89US-0382802.
XX
PA (CALJ) CALGENE INC.
XX
PI Houck CM, Pear JR, Martineau B, Hiatt W;

XX
DR WPI; 1991-024190/04.
XX
PT Modulating endogenous cytokinin levels - regulatory regions are
PT transformed into plant cells e.g. fruit to modify pheno-type
XX
PS Disclosure; Fig 3; 39pp; English.
XX
CC The sequence encodes an enzyme in the cytokinin biosynthetic pathway,
CC derived from Z130, and useful in modulating a transformed plant's
CC phenotype eg. fruit maturation, ripening etc.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;

Query Match   27.2%; Score 147.4; DB 12; Length 4383;
Best Local Similarity   67.1%; Pred. No. 2.5e-26;
Matches 261; Conservative   0; Mismatches 111; Indels 17; Gaps   3;

Qy   82 TGCAAAACAGAAGAGCAACACATTTCCTGGAATATGCATTACCAAACCACCATGCAGAAAA 141
         |||||| || || |||| ||||| |||| ||||| ||| || ||| |||||| ||||
Db 2942 TGCAAAGCACCAAGCCAAACTTTCCCAGGATTATGTTTTATGGACTCATCATGTAGAAAA 3001

Qy   142 GCTTGTATCAGTGAGAAAATTTACTGATGCTCATTGTTAGCAAAATCCTCAGAAGGTGCCTA 201
         || ||||||| ||||||||||||| | | |||||||||||||| ||| || |||||| ||||
Db 3002 TATTGTATCAAAGAGAAATTTACTCGGTGGACATTGTTAGCAAAACTCCAAAGAAGGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTGTTTGATGAGGAAGATGACTAAAAACAGGAGCTGAAAATTTTGGCT 261
         || ||||||||| |||| || ||| |||| ||| | ||| || ||| |||| |
Db 3062 TGCACTAAGCCAATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAAGAGAGATAATGGATTAACTAATTA 321
         ||||||||||||| | || | |||| ||||||||||||||||||| ||| |||| | ||
Db 3119 GAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAAGAAGAGAGATTATGTTTTATGGGGAGTAATAA 3178

Qy   322 GAGATTAGAAGAAAATTAAGGAT----------GCAGTATCACACATAATAAAAGTTTCTA 370
         ||| ||| ||||| |||||| ||||||| |||||||||||||||| | |
Db 3179 TTAAGTGAGGTTAAATAAGGAGATTTTGAGTGTCAAAAAAAACAAAAATTAAATAAAAGTGTTGC 3238

Qy   371 CCTTTCTTAAAAAGTGTAGC---TAAGTGTTGTTTTTAAATTGGCTTTTAGTTAGTTAGCCTTTTAT 427
         | ||||||| |||| ||| | || ||||| ||| ||| ||||| ||||||| ||
Db 3239 CTTTTCTTATTAGGGTGAGCTTGTGTCGTTGTTAGTTAGTTAGTAGCCATTGA 3298

Qy   428 TACACTTTAAATAAGTTTGTGTGGCACTTCAAT 456
         | |||| |||||||||||||| || | | ||
Db 3299 CACATTAAATAAGTTTGTGACACATCATT 3327

RESULT 14
AAQ35143

ID AAQ35143 standard; cDNA; 4383 BP.
XX
AC AAQ35143;
XX
DT 25-MAR-2003 (updated)
DT 10-MAR-2003 (updated)
DT 24-MAY-1993 (first entry)
XX
DE Calgene lambda 140/pz130 DNA including pz130 gene.
XX
KW pz130; expression; cytokinin; plant; ss.
XX
OS Lycopersicon esculentum.
OS Bacteriophage lambda.
XX
FH Key   Location/Qualifiers
FH
FT intron   2702..2921
FT   /*tag= a
FT misc_feature   2598..3382
FT   /*tag= b
FT   /note= "overlap between lambda 140 genomic clone
FT   with pz130 cDNA clone"
FT misc_RNA   2567

FT   /*tag= c
FT   /note= "start of pZ130 gene transcript"
XX
PN   US5177307-A.
XX
PD   05-JAN-1993.
XX
PF   17-JUL-1990;   90US-0554196.
XX
PR   26-MAY-1987;   87US-0054369.
PR   15-MAR-1988;   88US-0168190.
PR   29-APR-1988;   88US-0188361.
PR   19-JUL-1989;   89US-0382802.
PR   17-JUL-1990;   90US-0554196.
XX
PA   (CALJ) CALGENE INC.
XX
PI   Hiatt W, Houck CM, Martineau BM, Pear JR;
XX
DR   WPI; 1993-035826/04.
XX
PT   New plant expression cassettes - contg. DNA encoding an enzyme in
PT   a cytokinin metabolic pathway to modify cytokinin levels in
PT   plants
XX
PS   Disclosure; Fig 3; 29pp; English.
XX
CC   The sequence shows the complete sequence of the region of Calgene
CC   lambda 140 genomic clone that overlaps with the pZ130 cDNA clone and
CC   a partial sequence of regions 5', and 3' to that region. It is
CC   possible that small gaps in the sequence may exist upstream from the
CC   SalI site bound at position 808. The clone is used in the
CC   construction of a novel expression cassette comprising in the 5'-3'
CC   direction of transcription, as operably joined components; a DNA
CC   sequence encoding an enzyme in a cytokinin metabolic pathway and a
CC   transcription termination region. The enzyme is pref. DNA
CC   transferase. The expression cassette can be used to modify a plant
CC   phenotype by modifying cytokinin levels during a specific time and/or
CC   in specific tissues during the development of a plant. Control of
CC   cytokinin concns. can be used to increase the total solids content of
CC   ripe fruit, improve retention of early fruit and alter the time of
CC   onset of ripening.   See also AAQ35144-5.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 4383 BP; 1540 A; 628 C; 606 G; 1600 T; 9 other;

Query Match   27.2%; Score 147.4; DB 14; Length 4383;
Best Local Similarity   67.1%; Pred. No. 2.5e-26;
Matches 261; Conservative   0; Mismatches 111; Indels 17; Gaps   3;

Qy   82 TGCAAAACAGAAAGCAACACATTTCCTGGAATATGCATTACCAAACCACCATGCAGAAAA 141
            |||||||| || | || ||||||||||||||||||||||||||||||||||||||||||
Db   2942 TGCAAAAGCACCAAGCCAAACTTTCCCAGGAATATATGTTTTATGGACTCATCATGTAGAAAA 3001

Qy   142 GCTTGTATCAGTGAGAAAATTTACTTACTGGTCATGTTCAGTCAGAAAATCCTCAGAAGTGCCTA 201
            ||||| || |||| ||| || |||| ||||||| || ||||||||||||||||||| |||
Db   3002 TATTGTATCAAAGAGAAATTTACTGGTGACATTGTTAGCAAACTCCAAAGGAAGTGTCTA 3061

Qy   202 TGTACTAAGCCATGTGTCGTGTTGATGAAGATGACTAAAAACAGGAGCTCGAAAATTTTGGCT 261
            || |||| ||||| | ||||||||||||||||||| |||||||||| |||||||||||
Db   3062 TGCACTAAGCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTGGGT 3118

Qy   262 GAGGAAGCAAAAACTTTGGCGCTGCAGCTTGCTTGAAGAGAGAGAATAACTAAATTA 321
            |||||||||||| || | | | | ||| ||||| |||||||||||||||||||
Db   3119 GAGGAAGCAAAAACTCTAAGTGAAGTGTGCTTGAAGAGAGGAGTAATAA 3178

Qy   322 GAGATTAGAAGAAAATTAAGGAT---------------GCAGTATCACACATAAATAAAAGTTCTA 370
            ||| |||||| || ||| | | ||||||||||| | ||| |||||||
Db   3179 TTAAGTGAGAGTTAAATAAGGAGATTTTGAGTGTCAAAAAAACAAAAATTAATAAAAGTGTTGC 3238

Qy   371 CCTTTCTTAAAAGTGTAGC---TAAGTGTGTGTTTAATTGGCTTTTAGTAGCCTTTTAT 427
            || ||||||||||||| | ||| | ||||||||| ||||| |||||||||
Db   3239 CTTTTCTTAATTAGGGTAGGCTGGTTGTTGAGTGTTGGCACTTCAAT 3298

Db   3299 CACATTAAATAAGTTTGTGACACATCATT 3327

Qy   428 TACACTTTAAATAAGTGTGGCACTTCAAT 456

RESULT 15
AAQ34941
ID   AAQ34941 standard; DNA; 4383 BP.
XX
AC   AAQ34941;
XX
DT   25-MAR-2003 (updated)
DT   19-MAY-1993 (first entry)
XX
DE   Calgene Lambda 140 genomic clone.
XX
KW   cDNA; clone pZ130; anthesis; tomato; ovary; integumen; outer pericarp;
KW   fruit; developemnt; transcription; initiation; region; modulation;
KW   ovary-specific; endogenous; fruit product; exogenous; phenotype;
KW   lycopersicon esculentum; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
FT   misc_RNA   2599..3382
FT   /*tag= a
FT   /note= "interrupted pZ130 cDNA clone"
FT   intron   2702..2921
FT   /*tag= b
FT   misc_signal   2567
FT   /*tag= c
FT   /function= Start of pZ130 gene transcript
XX
PN   US5175095-A.
XX
PD   29-DEC-1992.
XX
DR   WPI; 1993-026940/03.
XX
PI   Houck CM, Martineau BM;
XX
PA   (CALJ) CALGENE INC.
XX
PT   DNA constructs contg. tomato pZ130 transcriptional initiation
PT   region - useful for modulation of endogenous fruit prods. and for
PT   prodn. of exogenous prods.
XX
PS   Disclosure; Fig 2; 18pp; English.
XX
CC   The sequence given shows the sequence of the region of the Calgene
CC   Lambda 140 genomic clone which overlaps with the pZ130 cDNA clone and
CC   a partial sequence 5', and 3' to that region. The pZ130 sequence is
CC   expressed during the early stages of anthesis in tomatoes. The message
CC   is expressed in ovary integumen and ovary outer pericarp tissue. It
CC   is not readily detectable in other tissues or at other stages of fruit
CC   development. The transcription initiation region associated with
CC   this gene is therefore considered to be ovary-specific. The actual
CC   function of the pZ130 polypeptide is unknown. The transcription
CC   initiation region can be used for modulation of endogenous fruit
CC   products, for production of exogenous products and for modification
CC   of the phenotype of fruit and fruit products.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;

```
Query Match        27.2%;   Score 147.4;  DB:14;   Length 4383;
Best Local Similarity   67.1%;   Pred. No. 2.5e-26;
Matches  261;  Conservative   0;  Mismatches 111;  Indels  17;  Gaps   3;

Qy      82 TGCAAAACAGAAAAGCAACACATTTCCTGGAATATGCATTACCAAACCACCATGCAGAAAA 141
           ||||||||||| ||  ||| ||  |||  |||| ||||| ||||| |||| || || ||||
Db    2942 TGCAAAAGCACCAAAGCCAAACTTTCCCAGGATTATGTTTATGGACTCATCATGTAGAAAA 3001

Qy     142 GCTTGTATCAGTGAGAAATTTACTGATGGGTCATTGTAGCAAAATCCTCAGAAGGTGCCTA 201
           ||||||| || | ||||| ||||| | |||||| |||||||| ||||| | |||| |||||
Db    3002 TATTGTATCAAAAGAGAAATTTACTGGTGGACATTGTAGCAAACTCCAAAGGAAAGTGTCTA 3061

Qy     202 TGTACTAAGCCATGTGTGTTTTGATGAAGATGACTAAAACAGGAGCTGAAATTTTTGGGCT 261
           | || |||||||| ||||| ||||| ||||  |||| |||||||| |  | |||| |||
Db    3062 TGCACTAAGCCCATGTGTATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTTTGGGT 3118

Qy     262 GAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAGAGATAATGGATAACTAATTA 321
           ||||| |||||||||| |||| ||| |||| |||||||| | || ||| || | |||
Db    3119 GAGGAAGCAAAAACTCTAAGTGAAGTTGTGCTTGAGAAGAGATTATGATGGAGTAATAA 3178

Qy     322 GAGATTAGAAGAGAAATTAAGGAT--------------GCAGTATCACACATAATAAAGTTTCTA 370
           |||||||  | || | |||||||              ||||||||||| |||||||||| ||||
Db    3179 TTAAGTGAGGTTAAATAAGGATTTTGAGTGGTCAAAAAAACAAAAATTAATAAAGTGTTGC 3238

Qy     371 CCTTTCTTTAAAAGTGTAGC---TAATGTTGTGTTTTAATTGGCTTTTAGTTAGCCTTTTAT 427
           ||||| | |||||||||||       ||||||| |||| || |||||||| ||||||||
Db    3239 CTTTTCTTTATTAGGGTAGCTTGTTGAGTTGTGTTTAGTTATTGGCCTATAGTTAGCCATTTGA 3298

Qy     428 TACACTTTAAATAAAGTGTGGGCACTTCAAT 456
           |||| |||||||||||||| |||||||| |
Db    3299 CACATTAAATAAGTTTGTGACACATCATT 3327
```

Search completed: January 28, 2004, 09:22:02
Job time : 224 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6

Run on: January 28, 2004, 08:46:51

Title: US-10-072-809B-18
Perfect score: 552
Sequence: 1 MARSLCFMAFAILARMLFVA..........LAEEAKTLAAALLEEEIMDN 105

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched: 1107863 seqs, 15872673 residues
Total number of hits satisfying chosen parameters: 1107863

Post-processing: Minimum Match 0%
                 Minimum DB seq length: 0
                 Maximum DB seq length: 200000000
                 listing first 45 summaries

Database : A_Geneseq_19Jun03:*
        1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
        2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
        3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
        4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
        5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
        6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
        7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
        8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
        9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
       10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
       11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
       12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
       13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
       14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
       15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
       16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
       17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
       18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
       19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
       20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
       21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
       22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
       23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
       24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	508	92.0	105	20	AAW83132	Nicotiana excelsio
2	468.5	84.9	106	20	AAW83133	Nicotiana panicula
3	342	62.0	79	20	AAW83131	Nicotiana excelsio
4	333	60.3	118	17	AAW97559	Tomato pZ130 cDNA-
5	333	60.3	118	18	AAW08364	Tomato ovary-speci
6	333	60.3	179	12	AAR10310	Ovary tissue trans
7	333	60.3	188	14	AAR30779	pZ130 polypeptide.
8	333	60.3	188	23	ABP56255	Plasmid pZ130.prot
9	197.5	35.8	84	22	AAB85079	Pepper defensin pr
10	190.5	34.5	84	23	ABU08325	Capsicum annuum L.
11	159.5	28.9	100	21	AAG07868	Arabidopsis thalia
12	159.5	28.9	128	21	AAG11734	Arabidopsis thalia
13	157.5	28.5	76	21	AAG25022	Arabidopsis thalia
14	149	27.0	55	21	AAG07870	Arabidopsis thalia
15	149	27.0	61	21	AAG07869	Arabidopsis thalia
16	149	27.0	61	21	AAG25023	Arabidopsis thalia
17	143.5	26.0	73	21	AAG07998	Arabidopsis thalia
18	143.5	26.0	73	21	AAG34807	Arabidopsis thalia
19	140.5	25.5	77	21	AAG09775	Arabidopsis thalia
20	138.5	25.1	120	21	AAB24880	Plant SDF encoded
21	138	25.0	58	21	AAB24880	Plant SDF encoded
22	138	25.0	58	21	AAG08000	Arabidopsis thalia
23	138	25.0	61	21	AAG34809	Arabidopsis thalia
24	138	25.0	61	21	AAG07999	Arabidopsis thalia
25	137.5	24.9	77	21	AAG04808	Arabidopsis thalia
26	137.5	24.9	77	21	AAB24660	Plant SDF encoded
27	137.5	24.9	77	21	AAB24882	Plant SDF encoded
28	137.5	24.9	77	21	AAG04820	Arabidopsis thalia
29	137.5	24.9	101	21	AAG38373	Arabidopsis thalia
30	137.5	24.9	109	21	AAB24659	Plant SDF encoded
31	137	24.8	78	23	AAB24881	Plant SDF encoded
32	137	24.8	78	23	AAU91018	Transplant media a
33	132.5	24.0	64	21	AAU91019	Transplant media a
34	132.5	24.0	64	21	AAG38375	Arabidopsis thalia
35	132.5	24.0	73	21	AAG38375	Arabidopsis thalia
36	132.5	24.0	73	21	AAG04821	Arabidopsis thalia
37	124	22.5	73	21	AAB85080	Pepper thionin-lik
38	121.5	22.0	62	21	AAB24661	Plant SDF encoded
39	121.5	22.0	62	21	AAG09776	Arabidopsis thalia
40	121	21.9	47	14	AAR33765	Potato gene pI322
41	120	21.7	75	17	AAR29752	Pepper Group 2 pro
42	120	21.7	75	22	AAE10676	Pepper group 2 pro
43	110	19.9	80	15	AAR33706	Antifungal protein
44	110	19.9	80	15	AAR57325	Antimicrobial Rs-A
45	110	19.9	80	18	AAW19280	Raphanus sativus a

RESULT 1
AAW83132

ID   AAW83132 standard; Protein; 105 AA.
XX
AC   AAW83132;
XX
DT   03-FEB-1999 (first entry)
XX
DE   Nicotiana excelsior thionine protein NeTHI2.
XX
KW   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
KW   NpTHI1; salt stress; resistance.
OS   Nicotiana excelsior.
XX
PN   JP10295360-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997; 97JP-0120179.
XX
PR   23-APR-1997; 97JP-0120179.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
DR   WPI; 1999-038278/04.
DR   N-PSDB; AAV70142.
XX
PT   Thionine gene derived by salt stress - used to deliver improved salt
PT   stress to plants
XX
PS   Claim 3; Page 5; 6pp; Japanese.

ALIGNMENTS

XX
CC   The present sequence represents a thionine protein from Nicotiana
CC   excelsior derived NeTHI2. The thionine protein has an effect of
CC   improving the salt stress resistance of a plant. The gene can improve
CC   the salt stress resistance of a plant.
XX
SQ   Sequence   105 AA;

Qy
Db

Qy   61 KILRRCLCTKPCVFDEKMIKTGAEILAEEAKTLAAALLEEEIMDN 105
        ||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKRPCVFDEKMIKTGAETLAEEATTLAAALLEEEIMDN 105

RESULT 2
AAW83133
ID   AAW83133 standard; Protein; 106 AA.
XX
AC   AAW83133;
XX
DT   03-FEB-1999 (first entry)
XX
DE   Nicotiana paniculata thionine protein NpTHI1.
XX
OS   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
KW   NpTHI1; salt stress; resistance.
OS   Nicotiana paniculata.
XX
PN   JP1029538O-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997; 97JP-0120179.
XX
PR   23-APR-1997; 97JP-0120179.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
DR   WPI; 1999-038278/04.
DR   N-PSDB; AAV70143.
XX
PT   Thionine gene derived by salt stress - used to deliver improved salt
PT   stress to plants
XX
PS   Claim 5; Page 5-6; 6pp; Japanese.
XX
CC   The present sequence represents a thionine protein from Nicotiana
CC   paniculata derived NpTHI1. The thionine protein has an effect of
CC   improving the salt stress resistance of a plant. The gene can improve
CC   the salt stress resistance of a plant.
XX
SQ   Sequence   106 AA;

Query Match   92.0%; Score 508; DB 20; Length 105;
Best Local Similarity   90.5%; Pred. No. 1.5e-45;
Matches   95; Conservative   5; Mismatches   5; Indels   0; Gaps   0;

Qy   1 MARSLCFMAFAILARMLFVAYEVQAREBCKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
         |||| |||||| ||||||||:|:|:||||||||||||||||||||||||||||||||||||
Db   1 MARSVCFMAFAILAVMLFVAYDVDEAKDCKTESNTFPGICITKPPCRKACISEKFTDGHCS 60

Qy   61 KILRRCLCTKPCVFDEKMIKTGAEILAEEAKTLAAALLEEEIMDN 105
        ||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKPCVFDEKMIKTGAEILAEEATTLAAALLEEEIMDN 106

Query Match   84.9%; Score 468.5; DB 20; Length 106;
Best Local Similarity   84.0%; Pred. No. 2e-41;
Matches   89; Conservative   8; Mismatches   8; Indels   1; Gaps   1;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
        ||||||||||| |||||||||||:   :||||||||||||||||||||||||||||
Db   1 MARSLCFMAFAVLAWMLFVAYEVQAKSTCKABSNTFPGLCITKPPCRKACLSEKFTDGKC 60

Qy   60 SKILRRCLCTKPCVFDEKMIKTGAEILAEEAKTLAAALLEEEIMDN 105
        ||||||:|||||||||||||||||||||:||||||||||||||:||
Db   61 SKILRRCICYKPCVFDGKMIQTGAENLAEEAETLAAALLEEEMMDN 106

RESULT 3
AAW83131
ID   AAW83131 standard; Protein; 79 AA.
XX
AC   AAW83131;
XX
DT   03-FEB-1999 (first entry)
XX
DE   Nicotiana excelsior thionine protein NeTHI1.
XX
OS   Nicotiana excelsior; Nicotiana paniculata; thionine; NeTHI1; NeTHI2;
KW   NpTHI1; salt stress; resistance.
OS   Nicotiana excelsior.
XX
PN   JP10295380-A.
XX
PD   10-NOV-1998.
XX
PF   23-APR-1997; 97JP-0120179.
XX
PR   23-APR-1997; 97JP-0120179.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
DR   WPI; 1999-038278/04.
DR   N-PSDB; AAV70141.
XX
PT   Thionine gene derived by salt stress - used to deliver improved salt
PT   stress to plants
XX
PS   Claim 1; Page 4; 6pp; Japanese.
XX
CC   The present sequence represents a thionine protein from Nicotiana
CC   excelsior derived NeTHI1. The thionine protein has an effect of
CC   improving the salt stress resistance of a plant. The gene can improve
CC   the salt stress resistance of a plant.
XX
SQ   Sequence   79 AA;

Query Match   62.0%; Score 342; DB 20; Length 79;
Best Local Similarity   77.5%; Pred. No. 2.4e-28;
Matches   69; Conservative   1; Mismatches   9; Indels   10; Gaps   2;

Qy   17 LFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPCVFDE 76
         ||||||||||:|:||||||||||||||||||||||||||||||||||||||||||
Db   1 LFVAYEVQARECARE--IFTGLCITNPQCRKACIKEKFTDGHCSKILRRCLCTKPC--- 54

Qy   77 KMIKTGAEILAEEAKTLAAALLEEEIMDN 105
        ||||||||||||||||||||||||||
Db   55 ----TGAETLAEEATTLAAALLEEEIMDN 79

RESULT 4
AAR97559
ID   AAR97559 standard; Protein; 118 AA.
XX
AC   AAR97559;
XX
DT   25-MAR-2003 (updated)
DT   14-SEP-1996 (first entry)
XX
DE   Tomato pZ130 cDNA-encoded protein.
XX
OS   Ovary; ovule; fruit; tomato; cotton; melanin; vector;
KW   transgenic plant.
OS   Lycopersicon esculentum cv. UC82B.
XX
PN   US5530185-A.
XX
PD   25-JUN-1996.

```
XX
PF   29-DEC-1992;   92US-0998158.
XX
PR   29-DEC-1992;   92US-0998158.
PR   19-JUL-1989;   89US-0382518.
PR   17-JUL-1990;   90US-0554195.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   Martineau BM, Reilley AA, Stalker DM;
XX
DR   N-PSDB; AAT31823.
XX
PT   DNA construct for expressing melanin synthesis gene in plant ovule
PT   cells - contains promoter from the tomato pz130 gene, also binary
PT   vector and transgenic plants, esp. cotton, contg. construct
XX
PS   Example 3; Fig 1A-B; 25pp; English.
XX
CC   The tomato pz130 clone contains a 564 bp insert of cDNA (AAT31823)
CC   detected only in a tomato cDNA library prepd. from pre-anthesis
CC   RNA. It codes for a protein (AAR97559) of unknown function. The
CC   clone was isolated from a library prepd. from cDNA of pre-anthesis
CC   stage ovaries. The pz130 transcriptional initiation region is
CC   considered to be ovary-specific and can be utilised in DNA
CC   constructs for the expression of heterologous genes, partic. in
CC   early fruit development.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   118 AA;

Query Match             60.3%;  Score 333;  DB 17;  Length 118;
Best Local Similarity   62.5%;  Pred. No. 3.3e-27;
Matches  65;  Conservative  12;  Mismatches  25;  Indels   2;  Gaps   2;

Qy          1  MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC  59
               |||:||||||||||:||| ||||:::|  |||||||||||||||| ||||||||||||||
Db         14  MARSIFFMAFLVLAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC  73

Qy         60  SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIM  103
               |:|||:||||||||| :::|||:||:|| ::|||||||||||||
Db         74  SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM  116


RESULT 5
AAW08364
ID   AAW08364 standard; Protein; 118 AA.
XX
AC   AAW08364;
XX
DT   14-MAR-1997  (first entry)
XX
DE   Tomato ovary-specific pz130-encoded protein.
XX
KW   Ovary; fruit; colour; pigmentation; cotton; tomato; probe;
KW   promoter; pz130; pz7; transgenic plant.
XX
OS   Lycopersicon esculentum cv. UC82B.
XX
PN   WO9640951-A2.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US09911.
XX
PR   07-JUN-1995;   95US-0487087.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   McBride K, Stalker DM;
XX
DR   WPI; 1997-052341/05.
DR   N-PSDB; AAT48813.
XX
PT   DNA construct capable of directing ovary-tissue transcription in
PT   plants - useful for modifying colour phenotype, in e.g. cotton
XX
PS   Example 1; Fig 1; 75pp; English.
XX
CC   Novel polypeptides (AAW08364 and AAW08365) are respectively encoded
CC   by tomato pz130 (AAT48813) and p270 (AAT48814) ovary-specific genes.
CC   Their native function is unknown. The promoter regions of the
CC   genes can be used to direct ovary-specific transcription of a
CC   gene of interest in transgenic plants.
XX
SQ   Sequence   118 AA;

Query Match             60.3%;  Score 333;  DB 18;  Length 118;
Best Local Similarity   62.5%;  Pred. No. 3.3e-27;
Matches  65;  Conservative  12;  Mismatches  25;  Indels   2;  Gaps   2;

Qy          1  MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC  59
               |||:||||||||||:||| ||||:::|  |||||||||||||||| ||||||||||||||
Db         14  MARSIFFMAFLVLAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC  73

Qy         60  SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIM  103
               |:|||:||||||||| :::|||:||:|| ::|||||||||||||
Db         74  SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM  116


RESULT 6
AAR10310
ID   AAR10310 standard; Protein; 179 AA.
XX
AC   AAR10310;
XX
DT   25-MAR-2003  (updated)
DT   04-APR-1991  (first entry)
XX
DE   Ovary tissue transcriptional factor DNA clone pz130 product.
XX
KW   Ovary tissue transcriptional factor; DNA construct; probe;
KW   clone pz130.
XX
OS   Lycopersicon esculentum UC82B.
XX
PN   EP409629-A.
XX
PD   23-JAN-1991.
XX
PF   19-JUL-1990;   90EP-0307926.
XX
PR   19-JUL-1989;   89US-0382518.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   Martineau B, Houck CM;
XX
DR   WPI; 1991-024191/04.
DR   N-PSDB; AAQ10262.
XX
PT   New ovary tissue transcriptional factors - modify transcription
PT   in tomato plant ovaries for use as mol. probes
XX
PS   Disclosure; Fig 1; 21pp; English.
XX
CC   The tomato-derived transcriptional initiation region which regulates
CC   the expression of the sequence corresp. to the pz130 clone is
CC   considered ovary-specific. Sequences hybridisable to the pz130
CC   clone, e.g. probe pz7, show abundant mRNA, esp. at the early stages
CC   of anthesis. The message is expressed in ovary integument and ovary
CC   outer pericarp tissue and is not expressed, or at any other stage
CC   detectable, in other tissues or at any other stage of fruit
CC   development. The native function of the amino acid sequence
CC   encoded by the structural gene comprising pz130 is unknown.
```

CC   See also AAQ10263-64.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   179 AA;

Query Match   60.3%; Score 333; DB 12; Length 179;
Best Local Similarity   62.5%; Pred. No. 5e-27;
Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
          |||:||||||||||||||||:|:|||| ||||||||||||||:|||||||||:|||||||
Db   14 MARSIFFMAFLVLAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFPTGGHC 73

Qy   60 SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIM 103
          |||:|:||:|||||||||:::|||||||||||||:||||||||
Db   74 SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM 116

RESULT 7
ID   AAR30779 standard; Protein; 188 AA.
AC   AAR30779;
XX
DT   25-MAR-2003 (updated)
DT   19-MAY-1993 (first entry)
DE   p2130 polypeptide.
XX
KW   cDNA; clone p2130; anthesis; tomato; ovary; integumen; outer pericarp;
KW   fruit; developement; transcription; initiation; region; modulation;
KW   ovary-specific; endogenous; fruit product; exogenous; phenotype.
OS   Lycopersicon esculentum.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 119
FT   /note= "Nonsense codon"
FT   Misc-difference 120
FT   /note= "Nonsense codon"
FT   Misc-difference 126
FT   /note= "Nonsense codon"
FT   Misc-difference 129
FT   /note= "Nonsense codon"
FT   Misc-difference 146
FT   /note= "Nonsense codon"
FT   Misc-difference 152
FT   /note= "Nonsense codon"
FT   Misc-difference 160
FT   /note= "Nonsense codon"
FT   Misc-difference 164
FT   /note= "Nonsense codon"
FT   Misc-difference 181
FT   /note= "Nonsense codon"

PN   US5175095-A.
XX
PD   29-DEC-1992.
XX
PF   17-JUL-1990; 90US-0554195.
XX
PR   19-JUL-1989; 89US-0382518.
PR   17-JUL-1990; 90US-0554195.
XX
PA   (CALJ) CALGENE INC.
XX
PI   Houck CM, Martineau BM;
XX
DR   WPI; 1993-026940/03.
DR   N-PSDB; AAQ34940.
XX
PT   DNA constructs contg. tomato p2130 transcriptional initiation
PT   region - useful for modulation of endogenous fruit prods. and for

PT   prodn. of exogenous prods.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence represents the polypeptide of cDNA clone p2130. This
CC   sequence is expressed during the early stages of anthesis in tomatos.
CC   The message is expressed in ovary integumen and ovary outer pericarp
CC   tissue. It is not readily detectable in other tissues or at other
CC   stages of fruit development. The transcription initiation region
CC   associated with this gene is therefore considered to be ovary-
CC   specific. The actual function of the p2130 polypeptide is unknown.
CC   The transcription initiation region can be used for modification of
CC   endogenous fruit products, for production of fruit and fruit products
CC   for modification of the phenotype of fruit and fruit products and
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   188 AA;

Query Match   60.3%; Score 333; DB 14; Length 188;
Best Local Similarity   62.5%; Pred. No. 5.3e-27;
Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
          |||:||||||||||||||||:|:|||| ||||||||||||||:|||||||||:|||||||
Db   14 MARSIFFMAFLVLAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFPTGGHC 73

Qy   60 SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIM 103
          |||:|:||:|||||||||:::|||||||||||||:||||||||
Db   74 SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM 116

RESULT 8
ID   ABP56255 standard; Protein; 188 AA.
AC   ABP56255;
XX
DT   28-MAR-2003 (first entry)
XX
DE   Plasmid p2130 protein sequence p27.
XX
KW   Transcription factor; ovary tissue.
OS   Unspecified.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 119
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 120
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 126
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 129
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 146
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 152
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 160
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 164
FT   /label= unknown
FT   /note= "encoded by a stop codon"
FT   Misc-difference 181
FT   /label= unknown
FT   /note= "encoded by a stop codon"

```
XX
PN   CN1189856-A.
XX
PD   05-AUG-1998.
XX
PF   07-JUN-1996;  96CN-0195170.
XX
PF   07-JUN-1995;  95US-0480087.
XX
PA   (CALJ ) CALGENE INC.
XX
DR   WPI; 2002-733392/80.
DR   N-PSDB; AB221970.
XX
PT   Use of transcriptional factors of ovary tissue -
XX
PS   Example 3; Fig 1A-B; 56pp; Chinese.
XX
CC   The present invention describes the use of transcriptional factors of
CC   ovary tissue. The present sequence represents a protein sequence
CC   which is used in an example from the present invention.
XX
SQ   Sequence  188 AA;

Query Match   60.3%;  Score 333;  DB 23;  Length 188;
Best Local Similarity   62.5%;  Pred. No. 5.3e-27;
Matches   65;  Conservative  12;  Mismatches  25;  Indels   2;  Gaps   2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARE-CKIESNTPFGICITKPPCRKACISEKFTDGHC  59
            ||||||||||||||||||||||||||  ||:|||||||||||||||:|||  ||||||||
Db  14  MARSIFMAFLVLIAMMLFVTYEVEAQQICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC  73

Qy  60  SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEIM  103
            ||:||||||||||||:|::            ||||:||||
Db  74  SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM  116


RESULT 9
AAB85079
ID   AAB85079 standard; Protein; 84 AA.
XX
AC   AAB85079;
XX
DT   22-AUG-2001  (first entry)
XX
DE   Pepper defensin protein (PepDef) sequence.
XX
KW   Pepper; defensin; PepDef; thionin-like protein; PepThi; transgenic;
KW   phytopathogen; antimicrobial.
XX
OS   Capsicum annuum.
XX
XX
PN   EP1101771-A1.
XX
PD   23-MAY-2001.
XX
PF   15-NOV-1999;  99EP-0309059.
PF   15-NOV-1999;  99EP-0309059.
XX
PR   15-NOV-1999;  99EP-0309059.
XX
PA   (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.
XX
PI   Oh B, Ko MK,  Shin B,  Chung CH;
XX
DR   WPI; 2001-357927/38.
DR   N-PSDB; AAF83987.
XX
PT   New pathogen induced genes (pepper defensin protein gene and pepper
PT   thionin-like protein gene) from Capsicum annuum, useful for producing
PT   transgenic plants with enhanced resistance against phytopathogens, e.g.
PT   fungi or nematode -
XX
PS   Claim 1; Page 11; 22pp; English.
```

```
XX
CC   The invention provides new isolated nucleic acid molecules encoding a
CC   pepper defensin protein (PepDef) and a pepper thionin-like protein
CC   (PepThi). The PepDef and PepThi genes are useful for producing transgenic
CC   plants that exhibit enhanced resistance against phytopathogens, e.g.
CC   fungi, bacteria, viruses, nematode, mycoplasma-like organisms, parasitic
CC   higher plants, flagellate protozoa or insects. The present sequence
CC   represents the pepper defensin protein (PepDef) sequence.
XX
SQ   Sequence  84 AA;

Query Match   35.8%;  Score 197.5;  DB 22;  Length 84;
Best Local Similarity   45.5%;  Pred. No. 3.3e-13;
Matches   45;  Conservative   9;  Mismatches  28;  Indels  17;  Gaps   2;

Qy   1  MARSLCFMAFAILARMLFVAYEVQARECKIESNTPFGICITKPPCRKACI-SEKFTDGHC  59
            |||| |:|:||  |||:||| ||  |:|| |||||||||||:|||  |||  ||||||||
Db   1  MARSIYFMAFLVLIAMTLFVAYGVQGXEICCKELTKPVKCGSSDPLCQKLCMEKEKYEDGHC  60

Qy  60  SKILRRCLCTKPCVFDEKMTKTGAEIIAEEAKTLAAALL  98
            ||| |:||||||||| ||:   ||||||||||:|||||
Db  61  FTILSKCLCMKRC-------------NAKTLATELL  83


RESULT 10
ABU08325
ID   ABU08325 standard; Protein; 84 AA.
XX
AC   ABU08325;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Capsicum annuum L. cv. Hanbyul thionine.
XX
KW   Thionine; plant disease resistance; plant.
XX
OS   Capsicum annuum L. cv. Hanbyul.
XX
FH   Key             Location/Qualifiers
FH
FT   Misc-difference  5
FT                    /note= "Encoded by ATT"
FT   Misc-difference  74
FT                    /label= Xaa
FT                    /note= "Xaa is given as "Met" in the specification,
FT                    and is encoded by AAT"
XX
PN   KR200202
```

```
XX
PN   KR20020024732-A.
XX
PD   01-APR-2002.
XX
PF   26-SEP-2000;  2000KR-0056518.
XX
PR   26-SEP-2000;  2000KR-0056518.
XX
PA   (KOCH-) KOREA CHUNGANG EDUCATIONAL FOUND.
XX
PI   Hwang BG,  Kim YJ,  Lee SC;
XX
DR   WPI; 2002-747901/81.
DR   N-PSDB; ABX12800.
XX
PT   Thionine gene of capsicum annuum l. cv. hanbyul and probing method of
PT   resistance for plant diseases -
XX
PS   Example 2; Fig 2; 14pp; Korean.
XX
CC   The present invention relates to the thionine gene of Capsicum
CC   annuum L. cv. Hanbyul, and a probing method of resistance for
CC   plant diseases, caused by Xanthomonas campestris subsp. vesicatoria,
CC   Colletotrichum coccodes, Colletotrichum gloeosporioides and the like.
CC   The presents sequence represents Capsicum annuum L. cv. Hanbyul
CC   thionine.
```

```
SQ   Sequence   84 AA;

Query Match
Best Local Similarity   44.4%;
Matches   44;  Conservative   9;  Mismatches   29;  Indels   17;  Gaps   2;

                    Score 190.5;  DB 23;  Length 84;
                    Pred. No. 1.8e-12;

Qy    1  MARSLCFMAFAILLARMLFVAYEVQARBCKTESNTFPGICITKPPCRKACI-SEKFTDGHC  59
            |||:| ||::|: :|||||| | |:  |:  |    :|| |    | |:    |||||
Db    1  MARSTFMAFLVLIAVTVFVAYGVQGKEICCKELTKPVKCSSDPLCQKLCMEKEKYEDGHC  60

Qy   60  SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALL  98
            |::|| || :|                    |:||| ::|
Db   61  FTILSKCLCMKRC-----------------XAKTLATELI  83
```

RESULT 11
AAG07868

```
ID   AAG07868 standard; Protein; 100 AA.
AC   AAG07868;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 5184.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999; 99US-0121825.
PR   05-MAR-1999; 99US-0123180.
PR   09-MAR-1999; 99US-0123548.
PR   23-MAR-1999; 99US-0125788.
PR   25-MAR-1999; 99US-0126264.
PR   29-MAR-1999; 99US-0126785.
PR   01-APR-1999; 99US-0127462.
PR   06-APR-1999; 99US-0128334.
PR   06-APR-1999; 99US-0128714.
PR   08-APR-1999; 99US-0128845.
PR   16-APR-1999; 99US-0129845.
PR   19-APR-1999; 99US-0130077.
PR   21-APR-1999; 99US-0130449.
PR   23-APR-1999; 99US-0130510.
PR   23-APR-1999; 99US-0130891.
PR   28-APR-1999; 99US-0131449.
PR   30-APR-1999; 99US-0132048.
PR   30-APR-1999; 99US-0132407.
PR   04-MAY-1999; 99US-0132484.
PR   05-MAY-1999; 99US-0132485.
PR   06-MAY-1999; 99US-0132486.
PR   06-MAY-1999; 99US-0132487.
PR   07-MAY-1999; 99US-0132863.
PR   11-MAY-1999; 99US-0134256.
PR   14-MAY-1999; 99US-0134218.
PR   14-MAY-1999; 99US-0134219.
PR   14-MAY-1999; 99US-0134221.
PR   14-MAY-1999; 99US-0134270.
PR   18-MAY-1999; 99US-0134768.
PR   19-MAY-1999; 99US-0134941.
PR   20-MAY-1999; 99US-0135124.
PR   21-MAY-1999; 99US-0135353.
PR   24-MAY-1999; 99US-0135629.
PR   25-MAY-1999; 99US-0136021.
PR   27-MAY-1999; 99US-0136392.
PR   28-MAY-1999; 99US-0136782.
PR   01-JUN-1999; 99US-0137222.
PR   03-JUN-1999; 99US-0137528.
PR   04-JUN-1999; 99US-0137502.
PR   07-JUN-1999; 99US-0137724.
PR   08-JUN-1999; 99US-0138094.
PR   10-JUN-1999; 99US-0138540.
PR   10-JUN-1999; 99US-0138847.
PR   14-JUN-1999; 99US-0139119.
PR   16-JUN-1999; 99US-0139452.
PR   16-JUN-1999; 99US-0139453.
PR   17-JUN-1999; 99US-0139492.
PR   18-JUN-1999; 99US-0139454.
PR   18-JUN-1999; 99US-0139456.
PR   18-JUN-1999; 99US-0139457.
PR   18-JUN-1999; 99US-0139458.
PR   18-JUN-1999; 99US-0139459.
PR   18-JUN-1999; 99US-0139460.
PR   18-JUN-1999; 99US-0139461.
PR   18-JUN-1999; 99US-0139462.
PR   18-JUN-1999; 99US-0139463.
PR   18-JUN-1999; 99US-0139750.
PR   21-JUN-1999; 99US-0139763.
PR   22-JUN-1999; 99US-0139817.
PR   22-JUN-1999; 99US-0139899.
PR   23-JUN-1999; 99US-0140353.
PR   24-JUN-1999; 99US-0140354.
PR   24-JUN-1999; 99US-0140695.
PR   28-JUN-1999; 99US-0140823.
PR   29-JUN-1999; 99US-0140991.
PR   30-JUN-1999; 99US-0141287.
PR   01-JUL-1999; 99US-0141842.
PR   02-JUL-1999; 99US-0142154.
PR   06-JUL-1999; 99US-0142055.
PR   06-JUL-1999; 99US-0142390.
PR   08-JUL-1999; 99US-0142803.
PR   09-JUL-1999; 99US-0142920.
PR   12-JUL-1999; 99US-0142977.
PR   14-JUL-1999; 99US-0143542.
PR   14-JUL-1999; 99US-0143624.
PR   15-JUL-1999; 99US-0144005.
PR   15-JUL-1999; 99US-0144085.
PR   16-JUL-1999; 99US-0144086.
PR   19-JUL-1999; 99US-0144325.
PR   19-JUL-1999; 99US-0144331.
PR   19-JUL-1999; 99US-0144332.
PR   19-JUL-1999; 99US-0144333.
PR   19-JUL-1999; 99US-0144334.
PR   19-JUL-1999; 99US-0144335.
PR   20-JUL-1999; 99US-0144352.
PR   20-JUL-1999; 99US-0144632.
PR   20-JUL-1999; 99US-0144884.
PR   21-JUL-1999; 99US-0144814.
PR   21-JUL-1999; 99US-0145086.
PR   21-JUL-1999; 99US-0145088.
PR   22-JUL-1999; 99US-0145085.
PR   22-JUL-1999; 99US-0145087.
PR   22-JUL-1999; 99US-0145089.
PR   23-JUL-1999; 99US-0145145.
PR   23-JUL-1999; 99US-0145192.
PR   23-JUL-1999; 99US-0145218.
PR   26-JUL-1999; 99US-0145224.
PR   27-JUL-1999; 99US-0145276.
PR   27-JUL-1999; 99US-0145913.
PR   28-JUL-1999; 99US-0145918.
PR   28-JUL-1999; 99US-0145919.
PR   02-AUG-1999; 99US-0145951.
PR   02-AUG-1999; 99US-0146386.
PR   02-AUG-1999; 99US-0146388.
PR   03-AUG-1999; 99US-0146389.
PR   03-AUG-1999; 99US-0147038.
PR   04-AUG-1999; 99US-0147204.
PR   04-AUG-1999; 99US-0147302.
PR   05-AUG-1999; 99US-0147192.
```

RESULT 12
AAG11734
ID   AAG11734 standard; Protein; 128 AA.
XX
AC   AAG11734;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 10566.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.

Best Local Similarity   44.3%; Pred. No. 3.7e-09;
Matches 31; Conservative 10; Mismatches 28; Indels 1; Gaps 1;

Qy   3 RSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKI 62
          ||| :||| : :|| :: |||| : | ||| || : :| ::| ::: |:| :|||
Db 32 RDSFFVLFMILATVM-GPVRVEARTCETSSNLFNGPCLSSSNCAAVCHNEGFSDGDCRGF 90

Qy 63 LRRCLCTKPC 72
          ||||||:|| :|
Db 91 RRRCLCTRPC 100

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160999.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 05-APR-1999; 99US-0128234.
PR 06-APR-1999; 99US-0128714.
PR 08-APR-1999; 99US-0129845.
PR 16-APR-1999; 99US-0130077.
PR 19-APR-1999; 99US-0130510.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142154.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144804.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 28-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0151065.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153363.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155596.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 22-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match   28.9%;   Score 159.5;   DB 21;   Length 128;
Best Local Similarity   44.3%;   Pred No. 4.8e-09;
Matches   31;   Conservative   10;   Mismatches   28;   Indels   1;   Gaps   1;

```
Qy        3  RSLCPMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKI   62
                | :|  | ||:|  |:|                 | |  ::|    ::| |||| :|
Db       60  RDSFFVLFMILATVM-GPVRVEARTCETSSNLFNGPCLSSSNCANVCHNEGFSDGDCRGF  118

Qy       63  LRCCLCTKPC   72
              || ||||| |
Db      119  RRRCLCTRPC  128
```

PF   25-FEB-2000; 2000EP-0301439.
XX
PD   06-SEP-2000.
XX
PN   EP1033405-A2.
XX
OS   Arabidopsis thaliana.
XX
KW   termination sequence.
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   Protein identification; signal transduction pathway; metabolic pathway;
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 28917.
XX
DT   17-OCT-2000 (first entry)
XX
AC   AAG25022;
ID   AAG25022 standard; Protein; 76 AA.
XX
RESULT 13
AAG25022

PR	25-FEB-1999; 99US-0121825.	PR	14-JUN-1999; 99US-0139119.
PR	05-MAR-1999; 99US-0123180.	PR	16-JUN-1999; 99US-0139452.
PR	09-MAR-1999; 99US-0123548.	PR	16-JUN-1999; 99US-0139453.
PR	23-MAR-1999; 99US-0125788.	PR	17-JUN-1999; 99US-0139492.
PR	25-MAR-1999; 99US-0126264.	PR	18-JUN-1999; 99US-0139454.
PR	29-MAR-1999; 99US-0126785.	PR	18-JUN-1999; 99US-0139455.
PR	01-APR-1999; 99US-0127462.	PR	18-JUN-1999; 99US-0139456.
PR	06-APR-1999; 99US-0128234.	PR	18-JUN-1999; 99US-0139457.
PR	08-APR-1999; 99US-0128714.	PR	18-JUN-1999; 99US-0139458.
PR	16-APR-1999; 99US-0129845.	PR	18-JUN-1999; 99US-0139459.
PR	19-APR-1999; 99US-0130077.	PR	18-JUN-1999; 99US-0139460.
PR	21-APR-1999; 99US-0130449.	PR	18-JUN-1999; 99US-0139461.
PR	23-APR-1999; 99US-0130510.	PR	18-JUN-1999; 99US-0139462.
PR	23-APR-1999; 99US-0130891.	PR	18-JUN-1999; 99US-0139463.
PR	28-APR-1999; 99US-0131449.	PR	18-JUN-1999; 99US-0139750.
PR	30-APR-1999; 99US-0132048.	PR	21-JUN-1999; 99US-0139763.
PR	04-MAY-1999; 99US-0132484.	PR	22-JUN-1999; 99US-0139817.
PR	05-MAY-1999; 99US-0132485.	PR	23-JUN-1999; 99US-0139889.
PR	06-MAY-1999; 99US-0132486.	PR	23-JUN-1999; 99US-0140353.
PR	07-MAY-1999; 99US-0132487.	PR	23-JUN-1999; 99US-0140354.
PR	07-MAY-1999; 99US-0132863.	PR	24-JUN-1999; 99US-0140695.
PR	11-MAY-1999; 99US-0134256.	PR	28-JUN-1999; 99US-0140823.
PR	14-MAY-1999; 99US-0134218.	PR	29-JUN-1999; 99US-0140991.
PR	14-MAY-1999; 99US-0134219.	PR	30-JUN-1999; 99US-0141287.
PR	14-MAY-1999; 99US-0134221.	PR	01-JUL-1999; 99US-0141842.
PR	18-MAY-1999; 99US-0134370.	PR	01-JUL-1999; 99US-0142154.
PR	18-MAY-1999; 99US-0134768.	PR	02-JUL-1999; 99US-0142055.
PR	19-MAY-1999; 99US-0134941.	PR	06-JUL-1999; 99US-0142390.
PR	20-MAY-1999; 99US-0135124.	PR	08-JUL-1999; 99US-0142803.
PR	21-MAY-1999; 99US-0135353.	PR	09-JUL-1999; 99US-0142920.
PR	24-MAY-1999; 99US-0135629.	PR	12-JUL-1999; 99US-0142977.
PR	25-MAY-1999; 99US-0136021.	PR	13-JUL-1999; 99US-0143542.
PR	27-MAY-1999; 99US-0136382.	PR	14-JUL-1999; 99US-0143624.
PR	28-MAY-1999; 99US-0136782.	PR	15-JUL-1999; 99US-0144005.
PR	01-JUN-1999; 99US-0137222.	PR	16-JUL-1999; 99US-0144005.
PR	03-JUN-1999; 99US-0137528.	PR	16-JUL-1999; 99US-0144086.
PR	04-JUN-1999; 99US-0137502.	PR	19-JUL-1999; 99US-0143325.
PR	07-JUN-1999; 99US-0137724.	PR	19-JUL-1999; 99US-0144331.
PR	08-JUN-1999; 99US-0138094.	PR	19-JUL-1999; 99US-0144332.
PR	09-JUN-1999; 99US-0138540.	PR	19-JUL-1999; 99US-0144333.
PR	10-JUN-1999; 99US-0138847.	PR	19-JUL-1999; 99US-0144334.
		PR	20-JUL-1999; 99US-0144335.
		PR	20-JUL-1999; 99US-0144352.
		PR	20-JUL-1999; 99US-0144632.
		PR	20-JUL-1999; 99US-0144884.
		PR	21-JUL-1999; 99US-0144814.
		PR	21-JUL-1999; 99US-0145086.
		PR	21-JUL-1999; 99US-0145088.
		PR	22-JUL-1999; 99US-0145087.
		PR	22-JUL-1999; 99US-0145089.
		PR	22-JUL-1999; 99US-0145192.
		PR	23-JUL-1999; 99US-0145145.
		PR	23-JUL-1999; 99US-0145218.
		PR	23-JUL-1999; 99US-0145224.
		PR	26-JUL-1999; 99US-0145276.
		PR	27-JUL-1999; 99US-0145913.
		PR	27-JUL-1999; 99US-0145918.
		PR	27-JUL-1999; 99US-0145919.
		PR	28-JUL-1999; 99US-0145951.
		PR	02-AUG-1999; 99US-0146386.
		PR	02-AUG-1999; 99US-0146388.
		PR	03-AUG-1999; 99US-0146389.
		PR	04-AUG-1999; 99US-0147038.
		PR	04-AUG-1999; 99US-0147204.
		PR	05-AUG-1999; 99US-0147302.
		PR	05-AUG-1999; 99US-0147192.
		PR	05-AUG-1999; 99US-0147260.
		PR	06-AUG-1999; 99US-0147303.
		PR	06-AUG-1999; 99US-0147416.
		PR	09-AUG-1999; 99US-0147493.
		PR	09-AUG-1999; 99US-0147935.
		PR	10-AUG-1999; 99US-0148171.

```
Qy   63 LRRCLCTKPC 72
           ||||:|||:|
Db   67 RRRCLCTRPC 76
```

RESULT 14

```
AAG07870
AAG07870;
AAG07870 standard; Protein; 55 AA.

AAG07870;

17-OCT-2000  (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 5186.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.
```

```
PR  25-FEB-1999;  99US-0121825.
PR  05-MAR-1999;  99US-0123180.
PR  09-MAR-1999;  99US-0123548.
PR  23-MAR-1999;  99US-0125788.
PR  25-MAR-1999;  99US-0126264.
PR  29-MAR-1999;  99US-0126785.
PR  01-APR-1999;  99US-0127462.
PR  01-APR-1999;  99US-0127465.
PR  06-APR-1999;  99US-0128234.
PR  08-APR-1999;  99US-0128714.
PR  08-APR-1999;  99US-0129845.
PR  16-APR-1999;  99US-0130077.
PR  19-APR-1999;  99US-0130449.
PR  21-APR-1999;  99US-0130510.
PR  23-APR-1999;  99US-0130891.
PR  28-APR-1999;  99US-0131449.
PR  30-APR-1999;  99US-0132048.
PR  30-APR-1999;  99US-0132407.
PR  04-MAY-1999;  99US-0132484.
PR  05-MAY-1999;  99US-0132485.
PR  06-MAY-1999;  99US-0132486.
PR  06-MAY-1999;  99US-0132487.
PR  07-MAY-1999;  99US-0132863.
PR  11-MAY-1999;  99US-0134256.
PR  14-MAY-1999;  99US-0134218.
PR  14-MAY-1999;  99US-0134219.
PR  14-MAY-1999;  99US-0134221.
PR  14-MAY-1999;  99US-0134370.
PR  18-MAY-1999;  99US-0134768.
PR  19-MAY-1999;  99US-0134941.
PR  20-MAY-1999;  99US-0135124.
PR  21-MAY-1999;  99US-0135353.
PR  24-MAY-1999;  99US-0135629.
PR  25-MAY-1999;  99US-0136021.
PR  27-MAY-1999;  99US-0136392.
PR  28-MAY-1999;  99US-0136782.
PR  01-JUN-1999;  99US-0137222.
PR  03-JUN-1999;  99US-0137528.
PR  04-JUN-1999;  99US-0137502.
PR  07-JUN-1999;  99US-0137724.
PR  08-JUN-1999;  99US-0138094.
PR  08-JUN-1999;  99US-0138540.
PR  10-JUN-1999;  99US-0138847.
PR  14-JUN-1999;  99US-0139119.
PR  16-JUN-1999;  99US-0139452.
PR  16-JUN-1999;  99US-0139453.
```

```
PR  11-AUG-1999;  99US-0148319.
PR  12-AUG-1999;  99US-0148341.
PR  13-AUG-1999;  99US-0148565.
PR  13-AUG-1999;  99US-0148684.
PR  16-AUG-1999;  99US-0149368.
PR  17-AUG-1999;  99US-0149175.
PR  18-AUG-1999;  99US-0149426.
PR  20-AUG-1999;  99US-0149722.
PR  20-AUG-1999;  99US-0149723.
PR  20-AUG-1999;  99US-0149929.
PR  23-AUG-1999;  99US-0149902.
PR  25-AUG-1999;  99US-0149930.
PR  25-AUG-1999;  99US-0150566.
PR  26-AUG-1999;  99US-0150584.
PR  27-AUG-1999;  99US-0151065.
PR  27-AUG-1999;  99US-0151066.
PR  27-AUG-1999;  99US-0151080.
PR  30-AUG-1999;  99US-0151303.
PR  31-AUG-1999;  99US-0151438.
PR  01-SEP-1999;  99US-0151930.
PR  07-SEP-1999;  99US-0152363.
PR  10-SEP-1999;  99US-0153070.
PR  13-SEP-1999;  99US-0153758.
PR  15-SEP-1999;  99US-0154018.
PR  16-SEP-1999;  99US-0154039.
PR  20-SEP-1999;  99US-0154779.
PR  22-SEP-1999;  99US-0155119.
PR  23-SEP-1999;  99US-0155486.
PR  24-SEP-1999;  99US-0155659.
PR  28-SEP-1999;  99US-0156458.
PR  29-SEP-1999;  99US-0156596.
PR  04-OCT-1999;  99US-0157117.
PR  05-OCT-1999;  99US-0157753.
PR  06-OCT-1999;  99US-0157865.
PR  08-OCT-1999;  99US-0158029.
PR  08-OCT-1999;  99US-0158232.
PR  12-OCT-1999;  99US-0158369.
PR  13-OCT-1999;  99US-0159293.
PR  13-OCT-1999;  99US-0159294.
PR  13-OCT-1999;  99US-0159295.
PR  14-OCT-1999;  99US-0159329.
PR  14-OCT-1999;  99US-0159330.
PR  14-OCT-1999;  99US-0159331.
PR  14-OCT-1999;  99US-0159637.
PR  14-OCT-1999;  99US-0159638.
PR  18-OCT-1999;  99US-0159584.
PR  21-OCT-1999;  99US-0160814.
PR  21-OCT-1999;  99US-0160815.
PR  21-OCT-1999;  99US-0160767.
PR  21-OCT-1999;  99US-0160768.
PR  21-OCT-1999;  99US-0160770.
PR  22-OCT-1999;  99US-0160980.
PR  22-OCT-1999;  99US-0160981.
PR  22-OCT-1999;  99US-0160989.
PR  25-OCT-1999;  99US-0161404.
PR  25-OCT-1999;  99US-0161405.
PR  25-OCT-1999;  99US-0161406.
PR  26-OCT-1999;  99US-0161359.
PR  26-OCT-1999;  99US-0161360.
PR  26-OCT-1999;  99US-0161361.
PR  27-OCT-1999;  99US-0161920.
PR  28-OCT-1999;  99US-0161992.
PR  28-OCT-1999;  99US-0161993.
PR  29-OCT-1999;  99US-0162142.
```

```
Query Match     28.5%;   Score 157.5;  DB 21;  Length 76;
Best Local Similarity   42.9%;  Pred. No. 4.5e-09;
Matches  30;  Conservative  11;  Mismatches  28;  Indels  1;  Gaps  1;

Qy   3 RSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKI 62
         :  :::: : :  :     :    ::   : :: :::  :  :::  :  ::: ::
Db   8 KSALLLFMILATVM-GPVRVEARTCETSSNLFNGPCLSSSNCANVCHNEGFSDGDCRGF 66
```

```
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   18-JUN-1999;   99US-0139817.
PR   21-JUN-1999;   99US-0139899.
PR   22-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   07-JUL-1999;   99US-0142803.
PR   08-JUL-1999;   99US-0142920.
PR   09-JUL-1999;   99US-0142977.
PR   12-JUL-1999;   99US-0143542.
PR   13-JUL-1999;   99US-0143624.
PR   14-JUL-1999;   99US-0144005.
PR   15-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144125.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144804.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   26-JUL-1999;   99US-0145224.
PR   27-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145513.
PR   28-JUL-1999;   99US-0145518.
PR   02-AUG-1999;   99US-0145919.
PR   02-AUG-1999;   99US-0145951.
PR   03-AUG-1999;   99US-0146386.
PR   04-AUG-1999;   99US-0146388.
PR   05-AUG-1999;   99US-0146389.
PR   05-AUG-1999;   99US-0147038.
PR   06-AUG-1999;   99US-0147204.
PR   09-AUG-1999;   99US-0147192.
PR   09-AUG-1999;   99US-0147260.
PR   10-AUG-1999;   99US-0147303.
PR   11-AUG-1999;   99US-0147416.
PR   12-AUG-1999;   99US-0147493.
PR   13-AUG-1999;   99US-0147935.
                    99US-0147171.
                    99US-0148319.
                    99US-0148341.
                    99US-0148565.

PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   23-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   30-AUG-1999;   99US-0151080.
PR   31-AUG-1999;   99US-0151303.
PR   01-SEP-1999;   99US-0151438.
PR   07-SEP-1999;   99US-0151930.
PR   10-SEP-1999;   99US-0152363.
PR   13-SEP-1999;   99US-0153070.
PR   15-SEP-1999;   99US-0153758.
PR   16-SEP-1999;   99US-0154018.
PR   19-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   18-OCT-1999;   99US-0159638.
PR   21-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   22-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   25-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   26-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   28-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   29-OCT-1999;   99US-0161993.
PR                  99US-0162142.
```

RESULT 15

```
Query Match        27.0%;    Score 149;  DB 21;  Length 55;
Best Local Similarity   50.0%;    Pred. No. 2.5e-08;
Matches   25;  Conservative   7;  Mismatches 18;  Indels 0;  Gaps 0;

Qy   23  VQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRCLCTKPC  72
             | || ||  | |  |    ||| | |      |||||||
Db    6  VEARTCETGSNLTNGPCLSSSNCANVCHNEGFSDGDCRGFRRCLCTRPC  55
```

```
AAG07869
ID   AAG07869 standard; Protein; 61 AA.
XX
AC   AAG07869;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 5185.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999; 99US-0121825.
PR   05-MAR-1999; 99US-0123180.
PR   09-MAR-1999; 99US-0123548.
PR   23-MAR-1999; 99US-0125788.
PR   25-MAR-1999; 99US-0126264.
PR   29-MAR-1999; 99US-0126785.
PR   01-APR-1999; 99US-0127462.
PR   06-APR-1999; 99US-0128234.
PR   08-APR-1999; 99US-0128714.
PR   16-APR-1999; 99US-0129845.
PR   19-APR-1999; 99US-0130077.
PR   21-APR-1999; 99US-0130449.
PR   23-APR-1999; 99US-0130510.
PR   28-APR-1999; 99US-0130891.
PR   30-APR-1999; 99US-0131449.
PR   30-APR-1999; 99US-0132048.
PR   04-MAY-1999; 99US-0132407.
PR   05-MAY-1999; 99US-0132484.
PR   05-MAY-1999; 99US-0132485.
PR   06-MAY-1999; 99US-0132486.
PR   06-MAY-1999; 99US-0132487.
PR   07-MAY-1999; 99US-0132863.
PR   11-MAY-1999; 99US-0134256.
PR   14-MAY-1999; 99US-0134218.
PR   14-MAY-1999; 99US-0134219.
PR   14-MAY-1999; 99US-0134221.
PR   14-MAY-1999; 99US-0134370.
PR   18-MAY-1999; 99US-0134768.
PR   19-MAY-1999; 99US-0134941.
PR   20-MAY-1999; 99US-0135124.
PR   21-MAY-1999; 99US-0135353.
PR   24-MAY-1999; 99US-0135629.
PR   25-MAY-1999; 99US-0136021.
PR   27-MAY-1999; 99US-0136392.
PR   28-MAY-1999; 99US-0136782.
PR   01-JUN-1999; 99US-0137222.
PR   03-JUN-1999; 99US-0137528.
PR   04-JUN-1999; 99US-0137502.
PR   07-JUN-1999; 99US-0137724.
PR   08-JUN-1999; 99US-0138094.
PR   10-JUN-1999; 99US-0138540.
PR   14-JUN-1999; 99US-0138847.
PR   16-JUN-1999; 99US-0139119.
PR   16-JUN-1999; 99US-0139452.
PR   16-JUN-1999; 99US-0139453.
PR   17-JUN-1999; 99US-0139492.
PR   18-JUN-1999; 99US-0139454.
PR   18-JUN-1999; 99US-0139455.
PR   18-JUN-1999; 99US-0139456.
PR   18-JUN-1999; 99US-0139457.
PR   18-JUN-1999; 99US-0139458.
PR   18-JUN-1999; 99US-0139459.
PR   18-JUN-1999; 99US-0139460.
PR   18-JUN-1999; 99US-0139461.
PR   18-JUN-1999; 99US-0139462.
PR   18-JUN-1999; 99US-0139463.
PR   18-JUN-1999; 99US-0139750.
PR   18-JUN-1999; 99US-0139763.
PR   21-JUN-1999; 99US-0139817.
PR   22-JUN-1999; 99US-0139899.
PR   23-JUN-1999; 99US-0140353.
PR   23-JUN-1999; 99US-0140354.
PR   24-JUN-1999; 99US-0140691.
PR   28-JUN-1999; 99US-0140823.
PR   29-JUN-1999; 99US-0140991.
PR   30-JUN-1999; 99US-0141287.
PR   01-JUL-1999; 99US-0141842.
PR   01-JUL-1999; 99US-0142154.
PR   02-JUL-1999; 99US-0142055.
PR   06-JUL-1999; 99US-0142390.
PR   08-JUL-1999; 99US-0142803.
PR   09-JUL-1999; 99US-0142920.
PR   12-JUL-1999; 99US-0142977.
PR   13-JUL-1999; 99US-0143542.
PR   14-JUL-1999; 99US-0143624.
PR   15-JUL-1999; 99US-0144005.
PR   16-JUL-1999; 99US-0144085.
PR   16-JUL-1999; 99US-0144086.
PR   19-JUL-1999; 99US-0144325.
PR   19-JUL-1999; 99US-0144331.
PR   19-JUL-1999; 99US-0144332.
PR   19-JUL-1999; 99US-0144333.
PR   19-JUL-1999; 99US-0144334.
PR   19-JUL-1999; 99US-0144335.
PR   20-JUL-1999; 99US-0144352.
PR   20-JUL-1999; 99US-0144632.
PR   20-JUL-1999; 99US-0144884.
PR   22-JUL-1999; 99US-0145085.
PR   22-JUL-1999; 99US-0145087.
PR   22-JUL-1999; 99US-0145089.
PR   22-JUL-1999; 99US-0145192.
PR   21-JUL-1999; 99US-0145145.
PR   23-JUL-1999; 99US-0145218.
PR   23-JUL-1999; 99US-0145224.
PR   26-JUL-1999; 99US-0145276.
PR   27-JUL-1999; 99US-0145913.
PR   27-JUL-1999; 99US-0145918.
PR   27-JUL-1999; 99US-0145919.
PR   28-JUL-1999; 99US-0145951.
PR   02-AUG-1999; 99US-0146386.
PR   02-AUG-1999; 99US-0146388.
PR   02-AUG-1999; 99US-0146389.
PR   03-AUG-1999; 99US-0147038.
PR   04-AUG-1999; 99US-0147204.
PR   04-AUG-1999; 99US-0147302.
PR   05-AUG-1999; 99US-0147192.
PR   05-AUG-1999; 99US-0147260.
PR   06-AUG-1999; 99US-0147303.
PR   06-AUG-1999; 99US-0147416.
PR   09-AUG-1999; 99US-0147493.
PR   09-AUG-1999; 99US-0147935.
PR   10-AUG-1999; 99US-0148171.
PR   11-AUG-1999; 99US-0148319.
PR   12-AUG-1999; 99US-0148341.
PR   13-AUG-1999; 99US-0148565.
PR   13-AUG-1999; 99US-0148684.
PR   16-AUG-1999; 99US-0193368.
PR   17-AUG-1999; 99US-0149175.
PR   18-AUG-1999; 99US-0149426.
PR   20-AUG-1999; 99US-0149722.
PR   20-AUG-1999; 99US-0149723.
PR   20-AUG-1999; 99US-0149929.
```

```
PR    23-AUG-1999;   99US-0149902.
PR    23-AUG-1999;   99US-0149930.
PR    25-AUG-1999;   99US-0150566.
PR    26-AUG-1999;   99US-0150884.
PR    27-AUG-1999;   99US-0151065.
PR    27-AUG-1999;   99US-0151066.
PR    27-AUG-1999;   99US-0151080.
PR    30-AUG-1999;   99US-0151303.
PR    31-AUG-1999;   99US-0151438.
PR    01-SEP-1999;   99US-0151930.
PR    07-SEP-1999;   99US-0152363.
PR    10-SEP-1999;   99US-0153070.
PR    13-SEP-1999;   99US-0153758.
PR    15-SEP-1999;   99US-0154018.
PR    16-SEP-1999;   99US-0154039.
PR    20-SEP-1999;   99US-0154779.
PR    22-SEP-1999;   99US-0155139.
PR    23-SEP-1999;   99US-0155486.
PR    24-SEP-1999;   99US-0155659.
PR    28-SEP-1999;   99US-0156458.
PR    29-SEP-1999;   99US-0156596.
PR    04-OCT-1999;   99US-0157117.
PR    05-OCT-1999;   99US-0157753.
PR    06-OCT-1999;   99US-0157865.
PR    07-OCT-1999;   99US-0158029.
PR    08-OCT-1999;   99US-0158232.
PR    12-OCT-1999;   99US-0158369.
PR    13-OCT-1999;   99US-0159293.
PR    13-OCT-1999;   99US-0159294.
PR    13-OCT-1999;   99US-0159295.
PR    14-OCT-1999;   99US-0159329.
PR    14-OCT-1999;   99US-0159330.
PR    14-OCT-1999;   99US-0159331.
PR    14-OCT-1999;   99US-0159584.
PR    18-OCT-1999;   99US-0159637.
PR    21-OCT-1999;   99US-0159638.
PR    21-OCT-1999;   99US-0160741.
PR    21-OCT-1999;   99US-0160767.
PR    21-OCT-1999;   99US-0160768.
PR    22-OCT-1999;   99US-0160770.
PR    22-OCT-1999;   99US-0160814.
PR    25-OCT-1999;   99US-0160815.
PR    25-OCT-1999;   99US-0160980.
PR    25-OCT-1999;   99US-0160981.
PR    25-OCT-1999;   99US-0160989.
PR    26-OCT-1999;   99US-0161360.
PR    26-OCT-1999;   99US-0161361.
PR    28-OCT-1999;   99US-0161404.
PR    28-OCT-1999;   99US-0161405.
PR    28-OCT-1999;   99US-0161406.
PR    29-OCT-1999;   99US-0161920.
PR    29-OCT-1999;   99US-0161992.
PR    29-OCT-1999;   99US-0161993.
PR    29-OCT-1999;   99US-0162142.
```

```
Query Match          27.0%;  Score 149;  DB 21;  Length 61;
Best Local Similarity 50.0%;  Pred. No. 2.8e-08;
Matches   25;  Conservative   7;  Mismatches  18;  Indels   0;  Gaps   0;

Qy   23 VQARECKYESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCCTKPC 72
        |.|| .|.|||.|.|::|.|:.||||||||..||||:|||.:|||:||.:||
Db   12 VEARTCETSSNLFNGPCLSSSNCANVCHNEGFSDGDCRGFRRRCLCTRPC 61
```

Search completed: January 28, 2004, 08:50:07
Job time : 42 secs

CM protein - protein search, using sw model.

Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 08:46:51 ; Search time 20 Seconds
(without alignments)
504.886 Million cell updates/sec

Title: US-10-072-809b-18
Perfect score: 552
Sequence: 1 MARSLCFMAFAILARMLFVA.........LAEEAKTLAAALLEEIMDN 105

Scoring table:

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
PIR_76:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

GenCore version 5.1.6

Scoring table: BLOSUM62
GapOp 10.0 , Gapext 0.5

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	529	95.8	105	2	S23574	thionin precursor,
2	333	60.3	105	2	S57809	gamma-thionin-like
3	150.5	27.3	79	2	T06381	proteinase inhibit
4	149	27.0	55	2	D96636	unknown protein, B
5	147.5	26.7	81	2	T02667	proteinase inhibit
6	144	26.1	78	2	S52634	gamma-thionin - Pe
7	143.5	26.0	73	2	D84433	proteinase inhibit
8	138.5	25.1	74	2	S05594	pseudothionin St1
9	137.5	24.9	77	2	S30578	proteinase inhibitor
10	137.5	24.9	77	2	C84433	proteinase inhibit
11	127.5	23.1	77	2	B84433	proteinase inhibit
11	125.5	22.7	77	2	T14395	proteinase inhibit
12	119	21.6	82	2	JC7897	defensin 1 precurs
13	117.5	21.3	75	2	S51637	probable proteinas
14	116	21.0	83	2	T14866	probable gamma-thi
15	112	20.3	80	2	T02622	probable antifunga
16	111	20.1	74	2	S24965	probable proteins
17	111	20.1	80	2	T10176	antifungal protein
18	108	19.5	80	2	T10183	antifungal protein
19	108	19.5	80	2	T10043	antifungal protein
20	107	19.4	47	2	A58319	gamma-zeathionin 1
21	106.5	19.3	79	2	T10043	antifungal protein
22	106	19.2	80	2	F96551	probable antifunga
23	106	19.2	80	2	T02621	probable antifunga
24	105	19.0	47	2	S57809	gamma-thionin SI-a
25	103	18.7	47	2	S69144	gamma-thionin SI-a
26	103	18.7	80	2	T10823	antifungal protein
27	103	18.7	80	2	F96787	antifungal protein
28	102.5	18.6	79	2	T07917	protein T4012.7 [i
29	101	18.3	47	2	A58445	fabatin-1 - fava b
30	101	18.3	47	2	B58445	fabatin-2 - fava b
31	99	17.9	47	2	S18849	hordothionin gamma
32	99	17.9	48	2	S13963	alpha-amylase inhi
33	94.5	17.1	74	2	T06766	disease resistance
34	94	17.0	82	2	S65779	gamma-thionin prec
35	92.5	16.8	75	2	S11530	purothionin gamma
36	88	15.9	72	2	S11156	pSAS10 protein - c
37	86.5	15.7	72	2	S11529	purothionin gamma
38	84.5	15.3	48	2	A55824	disease resistance
39	83	15.0	83	2	T06599	disease resistance
40	77	13.9	70	2	B58319	gamma-zeathionin 2
41	76	13.8	47	2	S20085	gamma-purothionin
42	73.5	13.3	297	2	T05673	omega-hordothionin
43	72.5	13.1	264	2	S29893	pin1 protein (clon
44	72	13.0	47	2	S69145	salivary glue prot
45	72	13.0	47	2	S69145	gamma-thionin SI-a

ALIGNMENTS

RESULT 1
S23574
thionin precursor, flower-specific - common tobacco
thionin precursor, flower-specific - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jun-2000
C;Accession: S23574; S21000
R;Gu, Q.; Kawata, E.E.; Morse, M.J.; Wu, H.M.; Cheung, A.Y.
Mol. Gen. Genet. 234, 89-96, 1992
A;Title: A flower-specific cDNA encoding a novel thionin in tobacco.
A;Reference number: S23574; MUID:92357021; PMID:1495489
A;Accession: S23574
A;Molecule type: mRNA
A;Residues: 1-105 <GTQ>
A;Cross-references: EMBL:Z11748; NID:g21212; PIDN:CAA77806.1; PID:g21213
A;Experimental source: strain W38; tissue type flower buds
A;Superfamily: gamma-thionin
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-105/Product: thionin #status predicted <MAT>
F;28-72,39-59,45-66,49-68/Disulfide bonds: #status predicted

Query Match 95.8%; Score 529; DB 2; Length 105;
Best Local Similarity 95.2%; Pred. No. 7.6e-48;
Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy   1 MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRXACISEKFTDGHCS 60
     |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1 MARSLCFMAFAILAWMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60

Qy   61 KILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEIMDN 105
      ||||||||||||||||||||||||||||||||||||||||||
Db   61 KILRRCLCTKPCVFDEKMIKTGAETLVEEAKTLAAALLEEIMDN 105

RESULT 2
S57809
gamma-thionin-like protein precursor - tomato
gamma-thionin-like protein precursor - tomato
N;Alternate names: AT2 protein; defensin
C;Species: Lycopersicon esculentum (tomato)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jun-2000
C;Accession: S57809; S72491
R;Milligan, S.B.; Gasser, C.S.
Plant Mol. Biol. 28, 691-711, 1995
A;Title: Nature and regulation of pistil-expressed genes in tomato.
A;Reference number: S57808; MUID:95375233; PMID:7647301
A;Accession: S57809
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-105 <MIL>
A;Cross-references: EMBL:U20591; NID:g924623; PIDN:AAA80496.1; PID:g924624
A;Experimental source: flower, cv. VF36
R;Brandstaedter, J.; Rossbach, C.; Theres, K.
Mol. Gen. Genet. 252, 146-154, 1996

A;Title: Expression of genes for a defensin and a proteinase inhibitor in specific areas
A;Reference number: S72491; MUID:96397493; PMID:8804387
A;Superfamily: gamma-thionin
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Reference number: 1-105 <BRA>
A;Experimental source: shoot, cv. Moneymaker
A;Accession: D96636
A;Status: preliminary
C;Genetics: DNA
A;Gene: AT2
C;Function:
A;Description: may respond to stimuli promoting the floral transition
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-74/Product: gamma-thionin-like protein #status predicted <MAT>
F;75-105/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;29-73,40-60,46-67,50-69/Disulfide bonds: #status predicted

Query Match   60.3%;   Score 333;   DB 2;   Length 105;
Best Local Similarity   62.5%;   Pred. No. 1.7e-27;
Matches   65;   Conservative   12;   Mismatches   25;   Indels   2;   Gaps   2;

Qy   1 MARSLCFMAFAILARMLFVAYEVQVARE-CKIESNTFPGICITKPPCRKACISEKFTDGHC 59
       |||||| ||| ||| || ||| || ||| |||||||||||||||||||||||||||||||
Db   1 MARSIFMAFLVLAMMLFVTYEVEAQOICKAPSQTFPGLCFMDSSCRKYCIKEKFTGGHC 60

Qy   60 SKILRRCLCTKPCVFDEKMTKTGAEILABEEAKTLAAALLEEIM 103
       |||||||||||||| ||:|
Db   61 SKLQRKCLCTKPCVFD-KISSEVKATLGEEAKTLSEVVLEEEIM 103

RESULT 3

T06381
proteinase inhibitor - soybean
C;Species: Glycine max (soybean)
C;Accession: T06381
R;Cushman, J. C.; Maitra, N.
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jun-2000
A;Description: Isolation and expression of a drought-induced protease inhibitor from soy
submitted to the EMBL Data Library, July 1994
A;Reference number: 215640
A;Accession: T06381
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-79 <CUS>
A;Cross-references: EMBL:U12150; NID:g533691; PIDN:AAC97524.1; PID:g533692
A;Experimental source: strain Essex; leaf
C;Superfamily: gamma-thionin

Query Match   27.3%;   Score 150.5;   DB 2;   Length 79;
Best Local Similarity   36.7%;   Pred. No. 1.2e-08;
Matches   29;   Conservative   16;   Mismatches   27;   Indels   7;   Gaps   1;

Qy   1 MARSLCFMAFAILARMLFVAYEV------QARECKTESNTFPGICITKPPCRKACISEK 53
       |:: : ::| | ||| :| |||:|| ||| || |||| | ||
Db   1 MSRSVPLVSTICVLLLLLVATEMMGPTMVAEARTCESQSHRFKGPCLSDTNCGSVCRTER 60

Qy   54 FTDGHCSKILRRCLCTKPC 72
       | |||| :|
Db   61 FTGGHCGRFRRCFCTKHC 79

RESULT 4

D96636
unknown protein, 87272-87105 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C;Accession: D96636
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
C;Authors: Hunter, J.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96636
A;Status: preliminary
C;Genetics: DNA
A;Gene: T7P1.20
C;Genetics:
A;Residues: 1-55 <STO>
A;Cross-references: GB:AE005173; NID:g6751695; PIDN:AAF27678.1; GSPDB:GN00141
A;Map position: 1
C;Superfamily: gamma-thionin

Query Match   27.0%;   Score 149;   DB 2;   Length 55;
Best Local Similarity   50.0%;   Pred. No. 1.2e-08;
Matches   25;   Conservative   7;   Mismatches   18;   Indels   0;   Gaps   0;

Qy   23 VQARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
       |||| ||| || ||| |:|| |: ||| |||
Db   6 VEARTCETSSNLFNGPCLSSSNCANVCHNEGFSGDGCRGFRRRCLCTKPC 55

RESULT 5

T02667
proteinase inhibitor - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jun-2000
C;Accession: T02667
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-81 <YUN>
R;Yun, C.H.; Lee, J.H.; Park, J.H.; Lee, G.R.
submitted to the EMBL Data Library, January 1998
A;Description: Molecular characterization of rice proteinase inhibitor gene.
A;Reference number: 214694
A;Accession: T02667
A;Cross-references: EMBL:AF044059; NID:g2829211; PIDN:AAC00503.1; PID:g2829212
A;Experimental source: strain IR36
C;Genetics:
A;Gene: RGPI9
A;Genetics:
A;Introns: 21/1
C;Superfamily: gamma-thionin

Query Match   26.7%;   Score 147.5;   DB 2;   Length 81;
Best Local Similarity   37.0%;   Pred. No. 2.4e-08;
Matches   27;   Conservative   13;   Mismatches   26;   Indels   7;   Gaps   1;

Qy   7 FMAFAILARMLFVAYEV------QARECKTESNTFPGICITKPPCRKACISEKFTDGHC 59
       : || :| | :: | |||:|| || || |||| | ||
Db   8 FSAMLLMVLLLAATGEMGGPVMVAEARTCESQSHRFKGPCARKANCASVCNTEGFPDGYC 67

Qy   60 SKILRRCLCTKPC 72
       ||||||| ||||
Db   68 HGVRRRCMCTKPC 80

RESULT 6

SS2634
gamma-thionin - Petunia inflata
C;Species: Petunia inflata
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jun-2000
C;Accession: SS2634
R;Karunanandaa, B.; Singh, A.; Kao, T.
Plant Mol. Biol. 26, 459-464, 1994
A;Title: Characterization of a predominantly pistil-expressed gene encoding a gamma-th
A;Reference number: SS2634; MUID:95036017; PMID:7948892
A;Accession: SS2634
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-78 <KAR>
A;Molecule type: DNA
A;Cross-references: GB:L27173; NID:g499654; PIDN:AAA64740.1; PID:g499655

C;Genetics:
A;Introns: 22/1
C;Superfamily: gamma-thionin

```
Query Match        26.1%;  Score 144;  DB 2;  Length 78;
Best Local Similarity  35.9%;  Pred. No. 5.4e-08;
Matches 28; Conservative 14; Mismatches 30; Indels  6; Gaps  1;

Qy   1 MARSLCFMAFAILARMLFVAYEV-------QARECKTESNTFPGICITKPPCPRKACISEKF 54
          |:|| ||:||: ||||| |      ||||||:|
Db   1 MGRSIRLFATFFLIAMLFLSTEMGPWTSAEARTCESQSHRFHGTCVRESNCASVCQTEGF 60

Qy  55 TDGHCSKILRCLCTKPC 72
          | |:|||||||||||
Db  61 IGGNCRAFRRCFCTRNC 78
```

RESULT 7
D84433
proteinase inhibitor II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C;Accession: D84433
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84433
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-73 <STO>
A;Cross-references: GB:AE002093; NID:g4038038; PIDN:AAC97220.1; GSPDB:GN00139
A;Gene: At2g02140
C;Genetics:
A;Gene: At2g02140
C;Superfamily: gamma-thionin

```
Query Match        26.0%;  Score 143.5; DB 2;  Length 73;
Best Local Similarity  40.0%;  Pred. No. 5.8e-08;
Matches 30; Conservative 13; Mismatches 27; Indels  5; Gaps  2;

Qy   1 MARSLCFMAFAILARMLFVAY-----EVQARECKTESNTFPGICITKPPCPRKACISEKFTDG 57
Db   1 MKLSLRLISALLMSVMLLFAYGMGPVEARTCESPSNKFQGVCLNSQSCAXACPSEGFSGG 60

Qy  58 HCSKILRCLCTKPC 72
Db  61 RCSSL--RCYCSKAC 73
```

RESULT 8
S05594
pseudothionin St1 precursor - potato (strain cv. Bintje)
C;Species: Solanum tuberosum (potato)
C;Variety: strain cv. Bintje
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jun-2000
C;Accession: S05594; S45659
R;Stiekema, W.J.; Heidekamp, F.; Dirkse, W.G.; van Beckum, J.; de Haan, P.; ten Bosch, C
Plant Mol. Biol. 11, 255-269, 1988
A;Title: Molecular cloning and analysis of four potato tuber mRNAs.
A;Reference number: S05594
A;Accession: S05592
A;Molecule type: mRNA
A;Residues: 1-74 <STI>
A;Cross-references: EMBL:X13180; NID:g21393; PIDN:CAA31577.1; PID:g21394
A;Experimental source: strain cv. Bintje
A;Note: it is unknown whether 1-Met is the initiator or whether translation is initiated
A;Note: the authors designated this protein as proteinase inhibitor (Bowman Birk) homolo
R;Moreno, M.; Segura, A.; Garcia-Olmedo, F.
Eur. J. Biochem. 223, 135-139, 1994
A;Title: Pseudothionin-St1, a potato peptide active against potato pathogens.

A;Reference number: S45659; MUID:94307252; PMID:8033886
A;Accession: S45659
A;Molecule type: protein
A;Residues: 28,'N',30-47 <MOR>
A;Experimental source: strain cv. Desiree
C;Superfamily: gamma-thionin
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-74/Product: pseudothionin St1 #status experimental <MAT>

```
Query Match        25.1%;  Score 138.5; DB 2;  Length 74;
Best Local Similarity  40.3%;  Pred. No. 1.9e-07;
Matches 29; Conservative 11; Mismatches 25; Indels  7; Gaps  2;

Qy   7 FMAFAILARMLFVAYEV-------QARECKTESNTFPGICITKPPCPRKACISEKFTDGHCS 60
Db   4 FATFFLIA-MLVVATKMGPWRIAEARHCSSLSHRFKGPCTRDSNCASVCETERFSGGNCH 62

Qy  61 KILRCLCTKPC 72
Db  63 GFRRCFCTKPC 74
```

RESULT 9
S30578
protease inhibitor II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Mar-2001
C;Accession: S30578; A84433
R;Yu, D.Y.; Quigley, F.; Mache, R.
submitted to the EMBL Data Library, November 1992
A;Description: Isolation and expression of a cDNA encoding protease inhibitor precursor
A;Reference number: S30579
A;Accession: S30578
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-77 <YUD>
A;Cross-references: EMBL:X69139; NID:g16426; PIDN:CAA48892.1; PID:g16427
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84433
C;Genetics:
A;Molecule type: DNA
A;Residues: 1-77 <STO>
A;Cross-references: GB:AE002093; NID:g4038041; PIDN:AAC97223.1; GSPDB:GN00139
A;Gene: At2g02100
A;Map position: 2
C;Superfamily: gamma-thionin

```
Query Match        24.9%;  Score 137.5; DB 2;  Length 77;
Best Local Similarity  33.8%;  Pred. No. 2.5e-07;
Matches 26; Conservative 16; Mismatches 30; Indels  5; Gaps  1;

Qy   1 MARSLCFMAFAILARMLFVA-------YEVQARECKTESNTFPGICITKPPCPRKACISEKFT 55
Db   1 MKLSMRLISAVLIMFMIFVATGMGPVTVEARTCESQSHRFKGTCVSASNCANVCHNEGFV 60

Qy  56 DGHCSKILRCLCTKPC 72
Db  61 GGNCRGFRRCFCTRHC 77
```

RESULT 10
C84433
proteinase inhibitor II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C;Accession: C84433
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Molecule type: DNA
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Accession: C84433
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-77 <STO>
A;Cross-references: GB;AE002093; NID:g4038039; PIDN:AAC97221.1; GSPDB:GN00139
C;Genetics:
A;Map Position: 2
A;Gene: At2g02130
C;Superfamily: gamma-thionin

```
Query Match           24.9%; Score 137.5; DB 2; Length 77;
Best Local Similarity 36.4%; Pred. No. 2.5e-07;
Matches   28; Conservative  15; Mismatches  29; Indels  5; Gaps  1;

Qy   1  MARSLCFMAFAILARMLFVA------YEVQARECKTESNTFPGICITKPPCRKACISEKFT  55
        :  |:|  :|  ||||
Db   1  MKLSVRFISAALLLFMVFIATGMGPVTVEARTCESKSHRFKGPCVSTHNCANVCHNEGFG  60

Qy  56  DGHCSKILRRCLCTKPC  72
        ||||:  :|  ||
Db  61  GGKCRGFRRRCYCTRHC  77
```

RESULT 11
B84433
proteinase inhibitor II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C;Accession: B84433
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Accession: B84433
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Molecule type: DNA
A;Residues: 1-77 <STO>
A;Cross-references: GB;AE002093; NID:g4038040; PIDN:AAC97222.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02120
C;Superfamily: gamma-thionin

```
Query Match           23.1%; Score 127.5; DB 2; Length 77;
Best Local Similarity 33.8%; Pred. No. 2.8e-06;
Matches   26; Conservative  12; Mismatches  34; Indels  5; Gaps  1;

Qy   1  MARSLCFMAFAILARMLFVA------YEVQARECKTESNTFPGICITKPPCRKACISEKFT  55
        :  |:|  :|  ||||
Db   1  MKFSMRIISAVLFLVMIFVATGMGPVTVEARTCASQSQRFKGKCVSDTNCENVCHNEGFP  60

Qy  56  DGHCSKILRRCLCTKPC  72
        ||||:  :|  ||
Db  61  GGDCRGFRRRCFCTRNC  77
```

RESULT 12
T14395
proteinase inhibitor II - turnip
C;Species: Brassica rapa (turnip)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C;Accession: T14395
R;Kim, H.U.; Yun, C.H.; Park, B.S.; Cho, W.S.; Ryu, J.C.; Chung, T.Y.
R;Kim, J.B.; Kim, ...; Lee, ...
submitted to the EMBL Data Library, March 1996
A;Description: Nucleotide sequences of protease inhibitor II in chinese cabbage.
A;Reference number: Z18029
A;Accession: T14395

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-77 <KIM>
A;Cross-references: EMBL;L31937; NID:g1209257; PID:g1209258
A;Experimental source: unopened flower bud
C;Superfamily: gamma-thionin

```
Query Match           22.7%; Score 125.5; DB 2; Length 77;
Best Local Similarity 33.8%; Pred. No. 4.5e-06;
Matches   26; Conservative  14; Mismatches  32; Indels  5; Gaps  1;

Qy   1  MARSLCFMAFAILARMLFVA------YEVQARECKTESNTFPGICITKPPCRKACISEKFT  55
        :  |:|  :|  ||||
Db   1  MKLSMRIISAVLLLFMIFVATGMGPVTVEARTCESKSHRFKGTCVSSTNCGNVCHNEGFG  60

Qy  56  DGHCSKILRRCLCTKPC  72
        ||||:  :|  ||
Db  61  GGKCRGFRVRCYCTRHC  77
```

RESULT 13
JC7897
defensin 1 precursor - wheat
C;Species: Triticum aestivum (common wheat)
C;Date: 03-Feb-2003 #sequence_revision 03-Feb-2003 #text_change 31-Mar-2003
C;Accession: JC7897
R;Koike, M.; Okamoto, T.; Tsuda, S.; Imai, R.
Biochem. Biophys. Res. Commun. 298, 46-53, 2002
A;Title: A novel plant defensin-like gene of winter wheat is specifically induced durin...
A;Reference number: JC7897; MUID:22266815; PMID:12379218
A;Accession: JC7897
A;Molecule type: mRNA
A;Residues: 1-82 <KOI>
A;Cross-references: DDBJ:AB089942
A;Experimental source: grown tissue
C;Comment: This protein is an antipathogenic protein and belongs to a subfamily of thio
d tolerance against pathogens during cold acclimation.
C;Genetics:
A;Gene: Tad1

```
Query Match           21.6%; Score 119; DB 2; Length 82;
Best Local Similarity 34.8%; Pred. No. 2.2e-05;
Matches   24; Conservative  13; Mismatches  24; Indels  8; Gaps  2;

Qy  12  ILARMLFVAYEV------QARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKIL--  63
        |:  |  |  :  |
Db  14  LLILLLIVATEMGTMKTAEARTCLSQSHKFKGTCLSNSNCAAVCRTENFPDGBCNTHLVE  73

Qy  64  RRLLCTKPC  72
        |:|  :|
Db  74  RKCYCKRTC  82
```

RESULT 14
S51637
probable proteinase inhibitor SE60 precursor - soybean
N;Alternate names: sulfur-rich protein SE60
C;Species: Glycine max (soybean)
C;Date: 01-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jun-2000
C;Accession: S51637; S51481
R;Choi, Y.; Ann, J.H.; Choi, Y.D.; Lee, J.S.
R;Lee, J.S.
Mol. Gen. Genet. 246, 266-268, 1995
A;Title: Tissue-specific and developmental regulation of a gene encoding a low molecula...
A;Reference number: S51481; MUID:95166185; PMID:7862099
A;Accession: S51481
submitted to the EMBL Data Library, July 1994
A;Reference number: S51637
A;Accession: S51637
A;Molecule type: DNA
A;Residues: 1-75 <LEE>
A;Cross-references: EMBL;Z18359; NID:g509768; PIDN:CAA79164.1; PID:g509769
A;Experimental source: strain Hwanggum

A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 29-75 <CHO>
A;Cross-references: GB:Z18359
A;Experimental source: strain Hwanggum
C;Genetics:
A;Introns: 21/1
C;Superfamily: gamma-thionin
C;Keywords: seed
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-75/Product: probable proteinase inhibitor SE60 #status predicted <MAT>

Query Match   21.8%; Score 117.5; DB 2; Length 75;
Best Local Similarity   33.3%;
Matches   25; Conservative   14; Mismatches   31; Indels   5; Gaps   2;

Qy   1 MARSLCFMAFAILARMLF---VAYEVQARBCKTESNTPGICITKPPCRKACISEKFTDG 57
            : ||:| |:| || :|| |:|||::|:|:|| :::::| :|:|| :|:|
Db   3 MRKSCGF--FFLLLILVFASQVVVQTEGRVCESQSHGFHGLCNRDHNCALVCRNEGFSGG 60

Qy   58 HCSKILRRCLCTKPC 72
            | : | |||:|:|:|
Db   61 RCKGFRRCPCTRIC 75

RESULT 15
T14866
probable gamma-thionin precursor SPI1 - Norway spruce
C;Species: Picea abies (Norway spruce)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C;Accession: T14866
R;Sharma, P.; Linneborg, A.
submitted to the EMBL Data Library, September 1995
A;Description: Isolation and characterization of a cDNA encoding a gamma-thionin-like pr
A;Reference number: Z18247
A;Accession: T14866
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-83 <SHA>
A;Cross-references: EMBL:X91487; NID:g1360107; PID:e198939
A;Experimental source: root
C;Genetics:
A;Gene: SPI1
C;Superfamily: gamma-thionin
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-83/Product: probable gamma-thionin precursor #status predicted <MAT>

Query Match   21.0%; Score 116; DB 2; Length 83;
Best Local Similarity   39.2%;
Matches   20; Conservative   8; Mismatches   21; Indels   2; Gaps   1;

Qy   24 QARECKTESNTPGICITKPPCRKACISEKFTDGHCSKIL--RRCLCTKPC 72
            :|||||:||||:||| |:|:::|:|:| :::| ::|:||: :: |:|:||
Db   32 EGRTCKTPSGKFKGVCASSNNCKNVCQTEGFPSGSCDFHVANRKCYCSKPC 82

Search completed: January 28, 2004, 08:48:40
Job time : 21 secs

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 08:44:46 ; Search time 1961 Seconds

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
(without alignments)
11286.134 Million cell updates/sec

Title: US-10-072-809B-17
Perfect score: 541
Sequence: 1 atggctcgctccttgtgctt..........tggttcaaaaaaaaaaaaa 541
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
    GenEmbl:*
    1: gb_ba:*
    2: gb_htg:*
    3: gb_in:*
    4: gb_om:*
    5: gb_ov:*
    6: gb_pat:*
    7: gb_ph:*
    8: gb_pl:*
    9: gb_pr:*
    10: gb_ro:*
    11: gb_sts:*
    12: gb_sy:*
    13: gb_un:*
    14: gb_vi:*
    15: em_ba:*
    16: em_fun:*
    17: em_hum:*
    18: em_in:*
    19: em_mu:*
    20: em_om:*
    21: em_or:*
    22: em_ov:*
    23: em_pat:*
    24: em_ph:*
    25: em_pl:*
    26: em_ro:*
    27: em_sts:*
    28: em_un:*
    29: em_vi:*
    30: em_vrt:*
    31: em_htg_hum:*
    32: em_htg_other:*
    33: em_htg_mus:*
    34: em_htg_pln:*
    35: em_htg_rod:*
    36: em_htg_mam:*
    37: em_htg_vrt:*
    38: em_sy:*
    39: em_htgo_hum:*
    40: em_htgo_mus:*
    41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	539.4	99.7	601	8	AF509566	AF509566 Nicotiana
	2	509.2	93.2	589	8	Z11748	Z11748 Solanaceae
	3	482.8	89.2	558	8	AB005266	AB005266 Nicotiana
	4	370.8	68.5	549	8	AB005250	AB005250 Nicotiana
	5	354.8	65.6	448	8	AB005265	AB005265 Nicotiana
	6	191.4	35.4	596	8	SLU20591	U20591 Solanum lyc
	7	163	30.1	559	8	AF128239	AF128239 Capsicum
	8	133.8	24.7	456	8	AF507975	AF507975 Capsicum
	9	122.2	22.6	602	8	AF507976	AF507976 Petunia x
	10	95.6	17.7	506	8	AF112443	AF112443 Capsicum
	11	95.6	17.7	548	8	AF112869	AF112869 Capsicum
	12	49.2	9.1	126368	9	HS1043E3	HS1043E3 Human DNA
c	13	49	9.1	172706	2	AL134847	AL134847 Human DNA
c	14	48.6	9.0	156060	2	AC004153	AC004153 Mus muscu
c	15	48.6	9.0	250707	3	AE014848	AE014848 Plasmodiu
	16	48.4	8.9	1335	1	AY182282	AY182282 Buchnera
	17	48	8.9	114694	9	AL355500	AL355500 Human DNA
	18	48	8.9	349980	6	AX344550	AX344550 Sequence
c	19	47.8	8.8	47296	3	AF160864	AF160864 Tetrahyme
	20	47.4	8.8	3361	4	AF482383	AF482383 Dictyoste
	21	47.4	8.8	15634	6	AX598816	AX598816 Sequence
	22	47.4	8.8	19634	6	AX598962	AX598962 Sequence
	23	46.2	8.5	1329	1	AY182262	AY182262 Buchnera
	24	46.2	8.5	1330	1	AY182253	AY182253 Buchnera
	25	46	8.5	57538	1	AC115682	AC115682 Dictyoste
	26	45.8	8.5	1335	1	AY182280	AY182280 Buchnera
	27	45.8	8.5	40780	3	CEM03C11	CEM03C11 Caenorhabdi
c	28	45.4	8.4	3016	3	DDU31631	DDU31631 Dictyosteli
c	29	45.4	8.4	4775	3	AF474338	AF474338 Dictyoste
c	30	45.4	8.4	7218	6	I66494	I66494 Sequence 14
c	31	45.4	8.4	232746	2	AC125778	AC125778 Rattus no
	32	45.4	8.4	271932	2	AC097673	AC097673 Rattus no
	33	45.4	8.4	272698	3	PFMAL4P4	PFMAL4P4 Plasmodiu
c	34	45.4	8.4	349600	1	AB063521	AB063521 Wiggleswo
	35	45.2	8.4	1335	1	AY182276	AY182276 Buchnera
	36	45.2	8.4	1335	1	AY182281	AY182281 Buchnera
	37	45.2	8.4	15363	3	MSQMTCG	MSQMTCG
	38	45.2	8.4	34548	6	AX349035	AX349035 Sequence
c	39	45.2	8.4	110000	3	AC116305	AC116305_0
c	40	45.2	8.4	167136	2	BX537323	BX537323 Danio rer
	41	45.2	8.4	266544	3	AC116956	AC116956 Dictyoste
	42	44.8	8.3	242927	2	BX248115	BX248115 Danio rer
	43	44.6	8.2	1301	1	AY182279	AY182279 Buchnera
	44	44.6	8.2	1330	1	AY182252	AY182252 Buchnera
	45	44.6	8.2	1330	1	AY182254	AY182254 Buchnera

ALIGNMENTS

RESULT 1
AF509566
LOCUS   AF509566   601 bp   mRNA   linear   PLN 19-MAR-2003
DEFINITION Nicotiana alata flower-specific defensin precursor (NaD1) mRNA,
            complete cds.
ACCESSION   AF509566
VERSION   AF509566.1 GI:25005096
KEYWORDS   .
SOURCE   Nicotiana alata (Persian tobacco)
ORGANISM   Nicotiana alata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1 (bases 1 to 601)
AUTHORS   Lay,F.T., Schirra,H.J., Scanlon,M.J., Anderson,M.A. and Craik,D.J.

```
  TITLE     The three-dimensional solution structure of NaD1, a new floral
            defensin from Nicotiana alata and its application to a homology
            model of the crop defense protein alfAFP
  JOURNAL   J. Mol. Biol. 325 (1), 175-188 (2003)
  MEDLINE   22361678
   PUBMED   12473460
REFERENCE   2  (bases 1 to 601)
  AUTHORS   Lay,F.T., Brugliera,F. and Anderson,M.A.
  TITLE     Isolation and Properties of Floral Defensins from Ornamental
            Tobacco and Petunia
  JOURNAL   Plant Physiol. 131 (3), 1283-1293 (2003)
   PUBMED   12644678
REFERENCE   3  (bases 1 to 601)
  AUTHORS   Lay,F.T. and Anderson,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2002) Biochemistry, La Trobe University,
            Bundoora, VIC 3086, Australia
FEATURES             Location/Qualifiers
     source          1..601
                     /organism="Nicotiana alata"
                     /mol_type="mRNA"
                     /db_xref="taxon:4087"
     gene            1..601
                     /gene="NaD1"
     CDS             61..378
                     /gene="NaD1"
                     /note="NaD1, gamma-thionin; possesses antifungal activity;
                     targeted to the vacuole"
                     /codon_start=1
                     /product="flower-specific defensin precursor"
                     /protein_id="AAN70999.1"
                     /db_xref="GI:25005097"
                     /translation="MARSLCFMAFAILAMMLFVAYEVQARECKTESNTFPGICITKPP
                     CRKACISEKFTDGHCSKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEI
                     MDN"
     mat_peptide     136..276
                     /gene="NaD1"
                     /product="flower-specific defensin"
BASE COUNT      192 a     95 c    115 g    199 t
ORIGIN

  Query Match             99.7%;  Score 539.4;  DB 8;  Length 601;
  Best Local Similarity   99.8%;  Pred. No. 6.3e-102;
  Matches  540;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         61 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAATGATGCTCTTTGTTGCC 120

Qy         61 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAGCAACACATTTCCTGGAATATGCATT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAGCAACACATTTCCTGGAATATGCATT 180

Qy        121 ACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 240

Qy        181 AAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGACTAAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGACTAAA 300

Qy        241 ACAGGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAGGAGCTGAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCAGCTTTGCTTGAAGAA 360

Qy        301 GAGATAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGATGCAGTATCACACATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGATAATGGATAACTAATTAGAGATTAGAAGAAATTAAGGATGCAGTATCACACATAAT 420

Qy        361 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTTGTGTTTTAATTGGCTTTTAGTAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGTTTCTACCTTTCTTAAAAGTGTAGCTAATGTTGTGTTTTAATTGGCTTTTAGTAGC 480
```

```
Qy        421 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAATCCTTTGTGCAATCTTGCACTAAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTTTATTACACTTTAAATAAGTGTGGCACTTCAATCCTTTGTGCAATCTTGCACTAAGT 540

Qy        481 TTATTTGTGTACTTTTAATGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTTGTGTACTTTTAATGAAAATGACCTTCTATGGTCTTTGGTTAAAAAAAAAAAAAA 600

Qy        541 A 541
              |
Db        601 A 601

RESULT 2
SFSTMRNA
LOCUS       SFSTMRNA                 589 bp    mRNA    linear   PLN 03-DEC-1993
DEFINITION  Solanaceae mRNA for flower-specific thionin.
ACCESSION   Z11748 S41631
VERSION     Z11748.1  GI:21212
KEYWORDS    thionin.
SOURCE      Solanaceae
  ORGANISM  Solanaceae
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Gu,Q., Kawata,E.E., Morse,M.J., Wu,H.M. and Cheung,A.Y.
  TITLE     A flower-specific cDNA encoding a novel thionin in tobacco
  JOURNAL   Mol. Gen. Genet. 234 (1), 89-96 (1992)
  MEDLINE   92357021
   PUBMED   1495489
REFERENCE   2  (bases 1 to 589)
  AUTHORS   Gu,Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-1992) Qing GU, Department of Biology, Yale
            University, 165 Prospect St, New Haven, CT, 06511, USA
FEATURES             Location/Qualifiers
     source          1..589
                     /organism="Solanaceae"
                     /mol_type="mRNA"
                     /db_xref="taxon:4070"
                     /clone="49A"
                     /tissue_type="flower of tobacco"
                     /dev_stage="flowering"
     CDS             49..366
                     /codon_start=1
                     /product="flower-specific thionin"
                     /protein_id="CAA77806.1"
                     /db_xref="GI:21213"
                     /db_xref="SWISS-PROT:P32026"
                     /translation="MARSLCFMAFAILAMMLFVAYEVQARECKTESNTFPGICITKPP
                     CRKACISEKFTDGHCSKLLRRCLCTKPCVFDEKMIKTGAETLVEEAKTLAAALLEEEI
                     MDN"
BASE COUNT      184 a     97 c    114 g    194 t
ORIGIN

  Query Match             93.2%;  Score 504.2;  DB 8;  Length 589;
  Best Local Similarity   95.7%;  Pred. No. 1.3e-94;
  Matches  518;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy          1 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAAGGATGCTCTTTGTTGCC 60
              ||||||||||||||||||||||||||||||||||||||||||| |||||| |||||||||
Db         49 ATGGCTCGCTCCTTGTGCTTCATGGCATTTGCTATCTTGGCAATGATGCTGTTTGTTGCC 108

Qy         61 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAGCAACACATTTCCTGGAATATGCATT 120
              |||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||
Db        109 TATGAGGTGCAAGCTAGAGAATGCAAAACAGAAAGCAATACATTCCCTGGAATATGCATT 168

Qy        121 ACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ACCAAACCACCATGCAGAAAAGCTTGTATCAGTGAGAAATTTACTGATGGTCATTGTAGC 228

Qy        181 AAAATCCTCAGAAGGTGCCTATGTACTAAGCCATGTGTGTTTGATGAGAAGATGACTAAA 240
```

RESULT 3

```
LOCUS       AB005266                558 bp    mRNA    linear   PLN 05-FEB-1999
DEFINITION  Nicotiana excelsior mRNA for gamma-thionin, complete cds.
ACCESSION   AB005266
VERSION     AB005266.1  GI:2244704
KEYWORDS    gamma-thionin.
SOURCE      Nicotiana excelsior
  ORGANISM  Nicotiana excelsior
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (sites)
  AUTHORS   Yamada,S., Komori,T. and Imaseki,H.
  TITLE     cDNA cloning of gamma-thionin from Nicotiana excelsior (Accession
            No. AB005266) (PGR97-131)
  JOURNAL   Plant Physiol. 115, 314 (1997)
REFERENCE   2  (bases 1 to 558)
  AUTHORS   Yamada,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-1997) Shigehiro Yamada, Plant Breeding and
            Genetics Research Laboratory, Japan Tobacco Inc.; 700 Higashibara,
            Iwata, Shizuoka 438, Japan
            (E-mail:Shigehiro.Yamada@pbgrl.jti.co.jp, Tel:0538-32-7116,
            Fax:0538-32-8700)
FEATURES             Location/Qualifiers
     source          1..558
                     /organism="Nicotiana excelsior"
                     /mol_type="mRNA"
                     /db_xref="taxon:61185"
     gene            1..558
                     /gene="NeThio2"
     CDS             33..350
                     /gene="NeThio2"
                     /codon_start=1
                     /product="gamma-thionin"
                     /protein_id="BAA21114.1"
                     /db_xref="GI:2244705"
                     /translation="MARSVCFMAFAILAVMLFVAYDVEAKDCKTESNTFPGICITKPP
                     CRKACIKEKFTDGHCSKILRRCLCTKPCVFDEKMIKTGAETLIAEEATLIAAALLEEI
                     MDN"

BASE COUNT      170 a     94 c    108 g    186 t
ORIGIN
```

Query Match   89.2%; Score 482.8; DB 8; Length 558;
Best Local Similarity   94.9%; Pred. No. 3.5e-90;
Matches 499; Conservative   0; Mismatches 27; Indels   0; Gaps   0;

RESULT 4

```
LOCUS       AB005250                549 bp    mRNA    linear   PLN 19-JUN-1999
DEFINITION  Nicotiana paniculata mRNA for gamma-thionin, complete cds.
ACCESSION   AB005250
VERSION     AB005250.1  GI:2251080
KEYWORDS    gamma-thionin.
SOURCE      Nicotiana paniculata
  ORGANISM  Nicotiana paniculata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (sites)
  AUTHORS   Komori,T., Yamada,S. and Imaseki,H.
  TITLE     A cDNA clone for gamma-thionin from Nicotiana paniculata (Accession
            No. AB005250) (PGR97-132)
  JOURNAL   Plant Physiol. 115, 314 (1997)
REFERENCE   2  (bases 1 to 549)
  AUTHORS   Komori,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-1997) Toshiyuki Komori, Plant Breeding and
            Genetics Research Laboratory, Japan Tobacco Inc.; 700 Higashibara,
            Iwata, Shizuoka 438, Japan
            (E-mail:Toshiyuki.Komori@pbgrl.jti.co.jp, Tel:0538-32-7116,
            Fax:0538-32-8700)
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Nicotiana paniculata"
                     /mol_type="mRNA"
                     /db_xref="taxon:62141"
     gene            1..549
                     /gene="NpThio1"
     CDS             48..368
```

Left column

```
FEATURES                Location/Qualifiers
     source              1..448
                         /organism="Nicotiana excelsior"
                         /mol_type="mRNA"
                         /db_xref="taxon:61185"
     gene                1..448
                         /gene="NeThio1"
     CDS                 <1..240
                         /gene="NeThio1"
                         /codon_start=1
                         /product="gamma-thionin"
                         /protein_id="BAA21113.1"
                         /db_xref="GI:2244703"
                         /translation="LFVAYEVQARECAREIFTGLCITNPQCRKACIKEKFTDGHCSKI
                         LRRCLCTKPCTGAETLAEEATTLAAALLEEEIMDN"
```

```
JOURNAL   Direct Submission
          Submitted (26-JUN-1997) Shigehiro Yamada, Plant Breeding and
          Genetics Research Laboratory, Japan Tobacco Inc.; 700 Higashibara,
          Iwata, Shizuoka 438, Japan
          (E-mail:Shigehiro.Yamada@pbgrl.jti.co.jp, Tel:0538-32-8700)
          Fax:0538-32-7116,
```

```
REFERENCE 2   (bases 1 to 448)
  AUTHORS   Yamada,S.
  TITLE     Direct Submission
  JOURNAL   Unpublished
```

```
REFERENCE 1   (sites)
  AUTHORS   Yamada,S., Komori,T. and Imaseki,H.
  TITLE     cDNA cloning of gamma-thionin from Nicotiana excelsior
  JOURNAL   Unpublished
```

```
ORGANISM  Nicotiana excelsior
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
SOURCE    Nicotiana excelsior
KEYWORDS  gamma-thionin.
VERSION   AB005265.1  GI:2244702
ACCESSION AB005265
DEFINITION Nicotiana excelsior mRNA for gamma-thionin, partial cds.
LOCUS     AB005265                  448 bp    mRNA    linear    PLN 04-JUL-1997
```

```
RESULT 5
AB005265
```

```
Qy          1 ATGGGCTCGCTCCTTGTGCTTCATCTTCTATCTTGGCAAGGAGATGCTCTTCGTTGTGCC 60
              ||||||| |||||||||||  || ||| |||||| ||||||||| |||| ||||||||||
Db         48 ATGGGCTCGCTCCTTGTGCTTCATCGTTCCGCAAGGAGATGCTCTTCGTTGTGCC 107
```

```
Qy         61 TATGAGGTGCAAGCT---AGAGAATGCAAAAACAGAAAGCAACACATTTCTGGAATATGC 117
              ||||| || ||||||   |||||||||||| ||||| |||||||| |||| |||||||||
Db        108 TATGAGGTGCAAGCTAAGAGAGAATGCAAAAACAGAAAGCAACACATTTCTGGAATATGC 167
```

```
Qy        118 ATTACCAAACCACCATGCAGAAAAAGCTTGTTCTCAGTGAGAAATTTACTGATGGTCATTGT 177
              |||| ||| ||| |||||||||||||||||||||||||||| |||||||| |||||||||
Db        168 ATTACCAAACCACCATGCAGAAAAAGCTTGTTCTCAGTGAGAAATTTACTGATGGTCATTGT 227
```

```
Qy        178 AGCAAAATCCTCAGAAGGTGCCTACTAGTACCTAAGCCATGGTGTATTTGATGAGGAAATGACT 237
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        228 AGCAAAATCCTCAGAAGGTGCATTTGCTACAAGCCATGTGTATTTGATGAAGAGATGATC 287
```

```
Qy        238 AAAACACGAGGAGCTGAAATTTTGGCTGAGGAAGCAACAAACTTTGCTTGAA 297
              ||||||||||||||||||||||| |||||||||||||||||||||
Db        288 CAAAACAGGAGCTGAAATTTGGGCCGAAGCAGAAGCAACTTGGCTTGAA 347
```

```
Qy        298 GAAGAGGATAATGACTAATACTAATTAGAGATTAGGAGAAATTAAGGAGCAGTATCACACAT 357
              |||||||| |||||| ||||||||| ||||||||| ||||||||||||||||||||||
Db        348 GAAGAGGATGGATAACTAATTAAGGAGGATGAAGTGTCACACAT 407
```

```
Qy        358 AATAAAGTTTCTACCTTTCTTCTTAAAGTAGTAGCTATAGTTGTGTGTGTTTAAGTAGT 417
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        408 AATAAAG-TGCTGCCTTTCTTAAAGTAGTAGCTAAGTGTGGTTTTTTTTAGT 466
```

```
Qy        418 AGCCTTTATTACACTTTAAATAAGTGTGTGGCGCACTTCAAATCCTTTGT 467
              ||||||||||||||||||||||||||||| |||||| || |||||||||
Db        467 AGCCGTTTGTTACACTTTTAAATAAGTGTGGGCACATCAAATCCTTTGT 512
```

Right column

```
FEATURES                Location/Qualifiers
     source              1..448
                         /organism="Nicotiana excelsior"
                         /mol_type="mRNA"
                         /db_xref="taxon:61185"
     gene                1..448
                         /gene="NeThio1"
     CDS                 <1..240
                         /gene="NeThio1"
                         /codon_start=1
                         /product="gamma-thionin"
                         /protein_id="BAA21125.1"
                         /db_xref="GI:2251081"
                         /translation="MARSICFMAFAVLAMMLFVAYEVQAKSTCKAESNTPQGICITKP
                         PCRKACLSEKFTDGKCSKILRRCICYKPCVFDGKMIQTGAENIAEEARTLAAALLEEE
                         MMDN"
```

```
Qy          1 ATGGCCCGCTCGCTGTGCTTCATGGCAGCAGTTGCTCTGGCCGTTGCCGTTGCC 60
              ||||||||||||||||  ||||| | || || ||| || |||||  ||||||||
Db         48 ATGGCTCCGCTTCGCTTGTGCTTCATGGCAGCAGTTGCTCTGGCCGTTGCC 107
```

```
Qy         61 TATGAGGTGCAAGCT---AGAGAATGCAAAAGCTCAAGCAACAGAAAGCAACACATTTATGC 117
              ||||| || ||||||   ||||||||| |||||||| ||||||||||||| |||||||||
Db        108 TATGAGGTGCAAGCTAAGAGAGAATGCAAAGCTCAAGCAACAGAAAGCAACACATTTATGC 167
```

```
Qy        168 ATTACCAAACCACCATGCAGAAAAAGCTTGTTCTCAGTGAGAAATTTACTGATGGTCATTGT 227
              ||||||||||||||||||||||||||||||||||||| |||||||||| |||||||||
Db        169 ATTACCAAACCACCATGCAGAAAAAGCTTGTTCTCAGTGAGAAATTTACTGATGGTCATTGT 237
```

```
RESULT 6
SLU20591
```

```
DEFINITION Solanum lycopersicum flower-specific gamma-thionin-like
           protein/acidic protein precursor, mRNA, complete cds.
ACCESSION  U20591
VERSION    U20591.1  GI:924623
KEYWORDS   .
SOURCE     Lycopersicon esculentum (tomato)
  ORGANISM Lycopersicon esculentum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Asteridae; lamiids; Solanales; Solanaceae; Lycopersicon.
REFERENCE  1  (bases 1 to 596)
  AUTHORS  Milligan,S.B. and Gasser,C.S.
  TITLE    Nature and regulation of pistil-expressed genes in tomato
  JOURNAL  Plant Mol. Biol. 28 (4), 691-711 (1995)
  MEDLINE  95375233
  PUBMED   7647301
```

```
Query Match          65.6%;  Score 354.8;  DB 8;  Length 448;
Best Local Similarity  88.1%;  Pred. No. 1.3e-63;
Matches  421;  Conservative  0;  Mismatches 27;  Indels 30;  Gaps  2;
```

```
Qy         49 CTCTTCGTTGTTGCCTATGAGGTGCAAGCTAGAGAATGCAGAAAGCAGAAGAATGCGCAAGAAAAACAGAAACAATTTCA------CT 108
              |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||  ||||
Db          1 CTCTTCGTTGTTGCCTATGAGGTGCAAGCTAGAGAATGCAGAAAGCAGAAGAATGCGCAAGAAAAACAGAAACAATTTCA------CT 54
```

```
Qy        109 GGAATATGCATTACCAAACCACCATGCAGAAAAAGCTTGTTCTGAGTGAGAAAGAAATTTACTGAT 168
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db         55 GGACTATGCATTACCAAATCCACAATGCAGAAAAAGCTTGTCATCAAAGAAATTTACTGAT 114
```

```
Qy        169 GGTCATTGTAGCAAAATCCTCAGAAGGTGCCTATGTTACTACTAAGCCATGTGTGTTTGATGAG 228
              ||||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||||
Db        115 GGTCATTGTAGCAAAATCCTCAGAAGGTGTCTATGCACTAAGCCATG---------- 161
```

```
Qy        229 AAGAGATGACTAAAACAGGAGCTGAAAATTTTGGGCTGCTGAGGAAGCAAAAAAACTTTGCTGCTGAGCT 288
              ||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        162 ---------CACAGGAGCTGAAAACTTTAGCTGAGGAAGCAACAACTTTGGCTGCTGAGCT 210
```

```
Qy        289 TTGCTTGAAGAAGAGAATGGGATAATGACTAATTAGGAGGATTAACTAATTAGAGAAATTAGGATGGCAGT 348
              ||||||||||||||||||||  ||||||| || |||||||||||  |||||||||||
Db        211 TTGCTTGAAGAAGAGAATGGGATAATTAACTAATTAGAGAATTAAGGATGGAGA 270
```

```
Qy        349 ATCACACATAATAAAAGTTTCTACCTTTCTTAAAAGTGTAGCTAAAGTGTGGTTTTTAAATTG 408
              ||||||||||||||||| ||||||||| ||||||||| |||||||||||||||||||
Db        271 GTCACACATAATAAAAGTTTCTACCTTTCTTAAAAGTGTAGCTAAAGTGTTGGTTTTAATTG 330
```

```
Qy        409 GCTTTTAGTAGCCCTTTTATTACACTTTAAATAAGTGTGTGGCACTTCAATCCTTCAATCCTTTGTGTGCAAT 468
              ||||||||||||||||  |||||||||||||||||||||||||  || ||||||||||||||
Db        331 GCTTTTAGTAGCCTTTTTATTACACTTTAAATAAGTGTGGCCACTTCAATCCTTTGTGTGCAAT 390
```

```
Qy        469 CTTGCACTAAGTTTTATTTGTGTACTTTTTTAAATGAAAAATGACCTTCTATGGTCTTTGGTT 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTTGCACTAAGTTTTATTCGTGTACTTTTTTAAATGAAAAATCACCTTCTATGGTTTTTGTTT 448
```

REFERENCE   2 (bases 1 to 596)
  AUTHORS   Milligan,S.B.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-FEB-1995) Stephen B. Milligan, Molecular and Cellular
            Biology, University of California at Davis, Davis, CA 95616, USA
FEATURES   Location/Qualifiers
     source   1..596
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /strain="VF36"
                     /db_xref="taxon:4081"
                     /clone="TPP3"
                     /tissue_type="pistil"
                     /clone_lib="lambda gt10 library from Gasser lab"
                     /dev_stage="immature"
     5'UTR   1..56
     CDS   57..374
                     /codon_start=1
                     /product="flower-specific gamma-thionin precursor"
                     /protein_id="AAB0496.1"
                     /db_xref="GI:924624"
                     /translation="MARSIFMAFLVLAMMLFVIVEVEAQQICTKAPSQTFPGLCFMDS
                     SCRKYCIKEKPTGGHCGSKLQRKCLCTKPCVPDKISSEVRATLGEEAKTLSEVVLEEEI
                     MME"
     sig_peptide   57..134
     mat_peptide   135..275
                     /product="gamma-thionin-like protein"
     mat_peptide   276..371
                     /product="acidic protein"
     3'UTR   375..596
BASE COUNT   204 a   82 c   107 g   203 t
ORIGIN

Query Match   35.4%; Score 191.4; DB 8; Length 596;
Best Local Similarity   66.2%; Pred. No. 9.5e-30;
Matches 370; Conservative   0; Mismatches 151; Indels   38; Gaps   5;

```
Qy       1 ATGGCTCGCTCCTTGTGCTTCATGGCATTGGCTATCCTTGGCAAGGATGGCTCTTGTTGCC 60
           ||||||||| ||| |   || ||||||||| || |  ||| |||| |||||||||| || |
Db      57 ATCGGCTCGTTCCATTTCTCTCAAGGCAATTTTGGTCTTGGCAATGATGGCTCTTTGTTACC 116

Qy      61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAAGCAACAACATTTCCTGGAATATGC 117
           |||||||||||   ||| ||||   ||||||  | | |||||||| || |||| |||| |||
Db     117 TATGAGGTGCAGAAGCTCAGCAAAATTGCAAAGCACCAAACTTCCCAAGGATAATGT 176

Qy     117 TATGAGGTGCAGAAGCTCAGCAAAATTTGCAAAGCACCAAACTTCCCAAGGATAATGT 176
           |||||||||||||| |  |||||     | ||||||| |   ||  | || ||||||
Db     177 TTTATGGACTCATCACGTAGAAAAATATTGTATCCAAAAGAAAAATTTACTCGGTGGACATTGT 236

Qy     118 ATTACCAAAACCACCATGCAGAAAAAGCTTGTATCAGTGAGAAATTTACTGAGTCATTGT 177
           ||| |  |||||| || |||| ||| ||||| ||| |||| |||||||  || || |||
Db     177 TTTATGGACTCATCACGTAGAAAAATATTGTATCCAAAAGAAAAATTTACTCGGTGGACATTGT 236

Qy     178 AGCAAAATCCTCAGAAGGACGTGCCTATTGTTACTAACGCCAATGTGTGTTTTGATGAAGAAT 237
           || |||||| || |||| | | ||| ||||  | ||| | |||  |||| |  |||| ||||
Db     237 AGCAAACTCCAAAAGGAGAGTGTCATCTAAGCCAAGCCATGTGTATT---TGACAAAATCTCA 293

Qy     238 AAAACAGGAGCTGAAAATTTTGGCTGAGGAAGAAGGAAAAAAACTCTAAGTGAAGTTGTGCTTGAA 297
           |||| | ||||||||||| ||| ||| | || ||| |||||| | |||| ||| || | || ||
Db     294 AGTGAAGTTAAAGCACACTTTGGGTGAGGAAGAAGGAAAAAAACTCTAAGTGAAGTTGTCGCTTGAA 353

Qy     298 GGAAGAGATAATGGATAAACTAATTAGAGAGAAATTAAGGAT---------GC 345
           |||||| |||| ||||| | | |||||| ||||| ||||||         ||
Db     354 GAAGAGATTATGATGGAGTAAATAATTAAGGAGTTAAATAAGGAAAAATTTACTGAGTGTCAAAA 413

Qy     346 AGTATCACACATAATAAAGTTCTACCTTTCTTAAAAGTGTAGC---TAATGTTGTGGTTT 402
           |||||||| |||| ||||| ||||| |||||  ||||||| ||   ||||| ||||| ||
Db     414 AAAAACAAAAATAATAAAGTGTTGCCTTTCTTATTATAGGGTAGGTTAATTAAGGTGTAGC 473

Qy     403 TAATTGGCTTTTAGTAGCCTTTTATTACACTTTAAATAAGTCGGCACTTCAATCCTTTG 462
           |||| ||| |||| || || | |||| |||  || |||||| || ||||  || | ||
Db     474 GTATTGGCCTATAGTAGCACAATTAAATAAGTTGTGACACATCATT--------- 527

Qy     463 TGCAATCTTGCACTAAGTTTATTTGTACTTTAATGAAAATGAAAATGACTTCTTATGTGTCTT 522
           |||||||||| |||| |||||||| || ||| |||||  ||| | |||||| ||||||||
Db     528 ---------------AATCCTTATGTATGTAGTTTTAATGAAAAAATGAGTCGACTAGCACGATGTT 576
```

RESULT 7

```
Qy     523 GGTTAAAAAAAAAAAAAAAAAAAA 541
           |||||||||||||
Db     577 AATTTTATATAAAAAAAAAAAAAA 595
```

LOCUS   AF128239   559 bp   mRNA   linear   PLN 06-JUL-1999
DEFINITION Capsicum chinense putative gamma-thionin precursor, mRNA, complete
            cds.
ACCESSION   AF128239
VERSION   AF128239.1 GI:4457222
SOURCE   Capsicum chinense
  ORGANISM Capsicum chinense
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1 (bases 1 to 559)
  AUTHORS   Aluru,M., Curry,J. and O'Connell,M.A.
  TITLE   Nucleotide sequence of a defensin or gamma-thionin-like gene
            (Accession No. AF128239) from habanero chile (PGR 99-070)
  JOURNAL   Plant Physiol. 120 (2), 633 (1999)
REFERENCE   2 (bases 1 to 559)
  AUTHORS   Curry,J., Aluru,M. and O'Connell,M.A.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-FEB-1999) Agronomy and Horticulture, New Mexico State
            University, P.O. Box 30003, MSC 3Q, Las Cruces, NM 88003, USA
FEATURES   Location/Qualifiers
     source   1..559
                     /organism="Capsicum chinense"
                     /mol_type="mRNA"
                     /cultivar="habanero"
                     /db_xref="taxon:80379"
     CDS   39..362
                     /note="defensin"
                     /codon_start=1
                     /product="putative gamma-thionin precursor"
                     /protein_id="AAD21200.1"
                     /db_xref="GI:4457223"
                     /translation="MARSIYFMAFLVLAVTLFVANGVQGQNNICKTTSKHFKGLCFAD
                     SKCRKVCIQEDKFEDGHCGSKLQRKCLCTKNCVPDNIPNDVGTILVQDAKTLEAQLLEE
                     EIGLI"
BASE COUNT   181 a   84 c   104 g   190 t
ORIGIN

Query Match   30.1%; Score 163; DB 8; Length 559;
Best Local Similarity   64.0%; Pred. No. 7.5e-24;
Matches 331; Conservative   0; Mismatches 170; Indels   16; Gaps   5;

```
Qy       1 ATGGCTCGCTCCTTGTGCTTCATGGCATTGGCTATCCTTGGCAAGGATGGCTCTTGTTGCC 60
           |||||||||| ||| |  || ||||| ||||| ||   || |||| | || ||||| |
Db      39 ATGGCTCGTTCCATTTACTTCATGGCATTTGCTAGTAACGTTGTTGCTAACGGTGTTCAGGGTCAAAATAACATTTGCT 98

Qy      61 TATGAGGTGCAAGCTAGAGA------ATGCAAAACAGAAAAGCAACAACATTTCCTGGAATA 114
           |||||||||||| | |  |       ||||||| |  || | | |  || || || |||||
Db      99 AATGGCTGGTCAAGGTCAAAATAACATTTGCAAAACAACAAGCAAGCACTTTAAAGGATTA 158

Qy     115 TGCATTACAAAACCACCATGCAGAAAAAGCTTGTATCAGTGAG---AAATTTACTGAGTGGT 171
           ||||| |||| |||   || ||||  | ||| |||  ||  ||   |||| |||| || |||
Db     159 TGTTTTGCCGACTCTAAATGTCGTAAAGTTTGTATTCAAGAAGATAAATTTGAAGATGGT 218

Qy     172 CATTGTTGAGCAAAATCCTCAGAAGGTGCCTATGTTGTTTGTTGATGAGAAG 231
           |||||| |||||| | |||   | || ||| || ||  | |||| | | ||
Db     219 CATTGTGGAAGCAAACTCCAAAAGGAAGTGCCTATGCACTAAGAACTGTGTT 275

Qy     232 ATGACTAAAAACAGGAGCTGAAAATTTTTGGCTGAGGAAGAAGGAAAAAAAACTTTGGCTGAGGAATA 291
           |  |||||  |  |  || ||| |||| |||||| |||||||||  |||| |  || | ||| ||
Db     276 AATCCTAATGATGTTGGAACTATTTTGGTTCAGGATGCTAAAAAACTCGAAGCACAATTG 335

Qy     292 CTTGAAGAAGAGATAATGGATAAACTAATTAGAGAGAAATTAAGGATCAGTAATC 351
           |||||||||||| |  | |||| ||  | ||||| | ||||| || |  || |||
Db                                                            
```

RESULT 8

```
LOCUS       AF507975                456 bp    mRNA            linear   PLN 19-MAR-2003
DEFINITION  Petunia x hybrida floral defensin-like protein 1 (D1) mRNA,
            complete cds.
ACCESSION   AF507975
VERSION     AF507975.1  GI:24817407
KEYWORDS    .
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Lay,F.T., Brugliera,F. and Anderson,M.A.
  TITLE     Isolation and Properties of Floral Defensins from Ornamental
            Tobacco and Petunia
  JOURNAL   Plant Physiol. 131 (3), 1283-1293 (2003)
  PUBMED    12644678
REFERENCE   2  (bases 1 to 456)
  AUTHORS   Brugliera,F., Holton,T.A., Stevenson,T. and Mason,J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2002) Florigene Ltd, 16 Gipps St., Collingwood,
            Melbourne, Victoria 3066, Australia
FEATURES             Location/Qualifiers
     source          1..456
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /cultivar="Old Glory Blue"
                     /db_xref="taxon:4102"
     gene            1..456
                     /gene="D1"
                     /note="PhD1"
     CDS             68..379
                     /gene="D1"
                     /codon_start=1
                     /product="floral defensin-like protein 1"
                     /protein_id="AAN64750.1"
                     /db_xref="GI:24817408"
                     /translation="MARSICFFAVIALAMLFAAYDAEAATCKAEACPTWDSVCINKKP
                     CVACCKKAKFSDGHCHSKILRRCLCTKECVPEKTEATQTETFTKDVNTLARLLEADMM
                     V"
BASE COUNT      124 a     92 c    102 g    138 t
ORIGIN
```

Query Match   24.7%; Score 133.8; DB 8; Length 456;
Best Local Similarity   65.0%; Pred. No. 9e-18;
Matches   214; Conservative   0; Mismatches   112; Indels   3; Gaps   1;

```
Qy      1 ATGGCTCGCTCCTTGTTGCTTCATCGTCTATCGTTGGCAAGGATGCTCTTTTGTTGCC 60
          ||||||||||||   | ||| ||||  ||| || || ||| |||||
Db     68 ATGGCTCGCTCCATCTGTTTCTTCGCCGTTGCTATACTAGCATTGGCA 127

Qy     61 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAAAAGCTTCTTTTGTTGCC 120
          |||||| ||| |||| ||||  ||||| |||| ||| ||| ||  ||||||
Db    128 TATGACGGAGCGGCCAACTTGCAAGGCAACACATTCCTGGAAATATGCATT 187

Qy    121 ACCAAACCACCATGCAGAAAAAGCTTGTATATCAGTGAGAAATTACTGATGGTTCATTGTGTAGC 180
          |||||||||||||||||||||||||||||
Db    336 CTTGAGAAGAAGAGAATTCCGGGCGTGTAATTAAGTCTGATTATGGATTTAGTTGTGT---CACAC 392
Qy    352 ACACATAATAAAGTTTCTACCTTTTCTTAAAAAGTGTAGCTAATGTGTTGTGTTTTAATTGGCT 411
          |||  |||       ||||     |||   | |||| |||   |||  ||  ||||||
Db    393 AAAATTAAATAAATGCGTTGCCTTTTCTTTAAAAGAGTAACTTACAATACTGTGTTCTTGGCT 452
Qy    412 TTTAGTAGCCTTTTATTACACTTTAAATAAGTGGGTGGCACTTCAAATCCTTTGTGCAATCTT 471
          |||||||||| |||| ||||| ||||| |||| ||| |||| |  |  || ||||| ||||||
Db    453 TATAGTAGCCATTTGACACA-TTAAATAAGTTGTGACAATAAATCCATATATGTGTTCTTG 511
Qy    472 GCACTAAGTTTTATTTGTGTTGTACTTTTTAATGAAAATGAC 508
          ||||  || ||||| |||||| ||  ||| ||||||| | |
Db    512 TATAAAGTTTTATGTGTTGTTTTAATGAAAAAGAGACCGTC 548
```

RESULT 9

```
LOCUS       AF507976                602 bp    mRNA            linear   PLN 19-MAR-2003
DEFINITION  Petunia x hybrida floral defensin-like protein 2 (D2) mRNA,
            complete cds.
ACCESSION   AF507976
VERSION     AF507976.1  GI:24817409
KEYWORDS    .
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Lay,F.T., Brugliera,F. and Anderson,M.A.
  TITLE     Isolation and Properties of Floral Defensins from Ornamental
            Tobacco and Petunia
  JOURNAL   Plant Physiol. 131 (3), 1283-1293 (2003)
  PUBMED    12644678
REFERENCE   2  (bases 1 to 602)
  AUTHORS   Brugliera,F., Herbert,S., Stevenson,T. and Mason,J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2002) Florigene Ltd, 16 Gipps St., Collingwood,
            Melbourne, Victoria 3066, Australia
FEATURES             Location/Qualifiers
     source          1..602
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /cultivar="Old Glory Blue"
                     /db_xref="taxon:4102"
     gene            1..602
                     /gene="D2"
                     /note="PhD2"
     CDS             67..372
                     /gene="D2"
                     /codon_start=1
                     /product="floral defensin-like protein 2"
                     /protein_id="AAN64751.1"
                     /db_xref="GI:24817410"
                     /translation="MARSICFFAVIALAMLFAAYETEAGTCKAECPTWEGICINKAP
                     CVKCCKAQPEKFTDGHCHSKILRRCLCTKPCATEEATATLANEVKTMAEALVEEDMME"
BASE COUNT      221 a    102 c    114 g    165 t
ORIGIN
```

Query Match   22.6%; Score 122.2; DB 8; Length 602;
Best Local Similarity   67.2%; Pred. No. 2.1e-15;
Matches   213; Conservative   0; Mismatches   83; Indels   21; Gaps   2;

```
Qy      1 ATGGCTCGCTCCTTGTTGCTTCATCGTCTATCGTTGGCAAGGATGCTCTTTTGTTGCC 60
          ||||||||||||   | ||| ||||  ||| || || ||| |||||
Db     67 ATGGCTCGCTCCATCTGTTTCTTCGCCGTTGCTATACTAGCATTGGCA 126

Qy     61 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAAAAGCTTCTTTTGTTGCC 120
          |||||| || |||||| ||||  ||||| |||| |||    ||  ||||||
Db    127 TATGAGACGACGAAGCGGGAAGCTTGGCAGGCAACACATTCCTGGAAGAATATGCATT 186

Qy    121 ACCAAACCACCATGCAGAAAAAGCTTGTATATCAGTGAGAAATTACTGATGGTTCATTGTGTAGC 180
          |||||||||||||||||||||||||||
Db    188 AATAAAAACCATCTGTAGCGTACGCGTTGCGGTGTTGCAAAAGGCAAAAGTTCTCGATGCTGAGC 247
Qy    181 AAAAATCCTCAGAGACGTGCCTATGTTGACTATCAAGCCATTGTCATTGGTTCTTTGTGAGGATGAC 240
          |  | |||  ||||    |  ||| ||| || |||||||| | |  || |||||
Db    248 AAAATACTCAGAGACGTGTCCACTAAGGACGTGTTAATTCAGAAAAAACTGAGGC---A 304
Qy    241 ACAGGAGCTGAAATTTGGCTGAGGAAGCAAAAAACTTTGGGCTGCGACGCTTTTGGTGAAGAA 300
          |||| |||||||     || || ||||||||   ||||||||| || ||| |||||  |||||
Db    305 ACTCAAACAGAAACTTTCACTAAGGATGTAAACACTTTGGCTCGACTTTTGGAGGCAA 364
Qy    301 GAGATAATGGATAACTAATTAGAGAATTAG 329
          |||||  || || ||  ||||||||||||
Db    365 GATATGATGGTGTAGAGGTTACTTGTTAG 393

Qy    121 ACCAAACCACCATGCAGAAAAGCTTGTA------TCAGTGAGAAATTACTGATGGTCAT 174
          |||||||||||||||||||||||||||      | |||||||||||||||||||||||
```

RESULT 10
LOCUS   AF112443   506 bp   mRNA   linear   PLN 10-DEC-1999
DEFINITION Capsicum annuum thionin-like protein (PepThi) mRNA, complete cds.
ACCESSION   AF112443
VERSION   AF112443.1 GI:6552501
KEYWORDS   .
SOURCE   Capsicum annuum
  ORGANISM Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1 (bases 1 to 506)
  AUTHORS   Oh,B.J., Ko,M.K., Kostenyuk,I., Shin,B. and Kim,K.S.
  TITLE   Coexpression of a defensin gene and a thionin-like gene via different
            signal transduction pathways in pepper and Colletotrichum
            gloeosporioides interactions
  JOURNAL   Plant Mol. Biol. 41 (3), 313-319 (1999)
  MEDLINE   20064969
   PUBMED   10598099
REFERENCE   2 (bases 1 to 506)
  AUTHORS   Oh,B.J., Ko,M.K., Kostenyuk,I.A., Shin,B.C. and Kim,K.S.
  TITLE   Direct Submission
  JOURNAL   Submitted (09-DEC-1998) Kumho Life and Environmental Science
            Laboratory, Kumho Petrochemical Co., LTD., 1 Oryong-dong, Puk-gu,
            Kwangju 500-712, Korea
FEATURES   Location/Qualifiers
     source   1..506
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4072"
     gene   1..506
                     /gene="PepThi"
     CDS   10..264
                     /gene="PepThi"
                     /codon_start=1
                     /product="thionin-like protein"
                     /protein_id="AAF16413.1"
                     /db_xref="GI:6552502"
                     /translation="MARSIYFMAFLVLAMTLFVAYGVQGKEICKELTKPVKCSSDPL
                     CQKLCMEKKYEDGHCFTILSKCLCMKRCNAKTLATELLA"

Db   187 AATAAACGCCCCATGTGTAAAAGTTGCAAAACACACCACCGACAAAAATTCACCGACGGCAC 246
         |||||| ||||||| ||||||||| |||||||||||||||||||||||||
Qy   175 TGTAGCAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATTGTGTTTGATGAGAAGATG 234
Db   247 TGCAGTAAAATACTTCCGAAGATGCCTATTGCACTAAGCCGTGTCAACTGAAGAG----- 300
         |||| |||||||| |||| ||||||||| ||||||| ||| ||||||
Qy   235 ACTAAAACAGGAGCTGAAAATTTTGGGCTGAGGAAGCAAAAACTTTGGGCTGCAGCTTTGCTT 294
Db   301 ---------GCAACTGCAACTTGGCCTAACGAGCTAAAGACTAATGGCGAAGCTTTGGGTC 351
                                 | | || |||||||||||||||| ||||||
Qy   295 GAAGAAGAGATAATGGA 311

Db   352 GAAGAAGATATGATGGA 368
         |||||||||||||||||
Qy

RESULT 11
LOCUS   AF112869   548 bp   mRNA   linear   PLN 19-DEC-1999
DEFINITION Capsicum annuum gamma-thionin 1 precursor (thion1) mRNA, complete
            cds.
ACCESSION   AF112869
VERSION   AF112869.1 GI:6601330
KEYWORDS   .
SOURCE   Capsicum annuum
  ORGANISM Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1 (bases 1 to 548)
  AUTHORS   Hwang,B.K., Lee,S.C., Kim,Y.J. and Hong,J.K.
  TITLE   Molecular cloning and pathogen-induced expression of a thionin gene
            in pepper plants
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 548)
  AUTHORS   Hwang,B.K., Lee,S.C., Kim,Y.J. and Hong,J.K.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-DEC-1998) Dept. of Agricultural Biology, Korea
            University, Anam-dong, Sungbuk-gu, Seoul 136-701, Korea
FEATURES   Location/Qualifiers
     source   1..548
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /cultivar="Hanbyul"
                     /db_xref="taxon:4072"
                     /tissue_type="leaf"
     gene   1..548
                     /gene="thion1"
     CDS   52..306
                     /gene="thion1"
                     /codon_start=1
                     /product="gamma-thionin 1 precursor"
                     /protein_id="AAF19936.1"
                     /db_xref="GI:6601331"
                     /translation="MARSIYFMAFLVLAMTVFVAYGVQGKEICCKELTKPVKCSSDPL
                     CQKLCMEKKYEDGHCFTILSKCLCMKRCNAKTLATELLA"

Db   130 AGTGACCTCTATGTCAAAAACTTGTATGGAGAAGGAGAAAATAATGAGAGATGCATTGT 189
         ||||||| ||||||||||||| | ||||||||||||||||||||| |||||||||
Qy   178 AGCAAAAATCCTCAGAAGGTGCCTATGTACTAAGCCATTGTGTTTGATGAGAAGATGACT 237
Db   190 TTCACAATCCTAAGCAAGTGCTTATGCATGAGAAGGAGATGTA------------------ 229
         | |||| || || | || ||||||||| ||
Qy   238 AAAACAGGAGCTGAAAATTTTGGGCTGAGGAAGCAAAAACTTTGGGCTGCAGCTTTGCTTGAA 297
Db   230 ------------------ATGCTAAAACTCTCGCAACAGAATTGCTTGCT 261
                                   | || | | ||| ||||||| ||
Qy   298 GAAGAGATAATGGATAACTAACTTTCAGATTAAGAGAAAATTTAAGGATGACTATCACACAT 357
Db   262 TGA---AGAAGAGTTCCCTCAAGCAGTAATTAAGTTTGATTATGAGATTAGTGTGTCACACAA 318
         || ||||||| |||||||||||||||||| |||||| | |||||||||||||
Qy   358 AATAAAGT---TTCTACCCTTTCTTAAAGAAGTAGCTAAGTGTTGTTTGATTAATGGCTTT 413
Db   319 AATTAAAATAAAGTGTTGCCCTTTCTTAAAAGGGTAACTTATAATTCTAGTTTCTTGGTA 378
         |||||||||||||| || |||||||| | | | |||| ||| || |||
Qy   414 TAGTAGGCCTTTTATTACACTTTAAATAAGTGTGGCACTTGTTGTCAATCTGC 473
Db   379 TAGTAGGCCATTGACACATTAAATTAAAATAAGTTGACACACATCAATCCTTCATTGTGTTCT 438
         ||||||||| || ||||| ||| ||||||| || ||||| ||| ||| ||| ||| ||
Qy   474 ACTAAGTTATTTGTGTACTTTTAAATGAAAATGACCTTCTATGGTCTTTGTTGGTTTAAA 533
Db   439 ATTAAG---TTTGTCTGTTTTAAATGAAAAGAGATCGTTTACGGGTCTAAAAAAAAAAAAAA 494
         ||||| |||| |||||||| | |||| |||| || ||| |||||||||||||||
Qy   534 AAAAAAAA 541
Db   495 AAAAAAAA 502

BASE COUNT   173 a   75 c   96 g   162 t
ORIGIN

Query Match   17.7%; Score 95.6; DB 8; Length 506;
Best Local Similarity   56.9%; Pred. No. 7.3e-10;
Matches 312; Conservative   0; Mismatches 174; Indels 62; Gaps   5;

Qy   1 ATGGCTCGCTCCTTGTTGCTTCATGCAATTGCTATATCTTGGGCAAGGGAGCTCTTTGTTTGCC 60
         ||||||||| ||||| |||||||||| ||||||| || ||||||||| ||||||||||||||
Db   10 ATGGCTCGTTCCATTACTTACTTCATGGCAAGGAGTGCTCATTGTTGTTTGTTGGC 69
Qy   61 TATGAGGTGCAAGGCAGTAGGAATGCAAAAACAGAAGCAACACATTTCCTGGGAATATGCAATT 120
         |||| ||||||||| | |||| | |||||||||||| ||||| | | | |||| | |
Db   70 TATGGGGTGCAAGGCAAGGAAATTTGCTGTAAAGAGCTCACAAAAACCTGTTAAATGTTCT 129
Qy   121 ACCAAACCACCATGCGAGAAAACGCTTGTAT---CAGTGAGAAATTACTGATGGTCATTGT 177
         || | |||||||||| |||||||| ||||||||| | |||| ||||
sig_peptide   52..114
                     CQKLCMEKKYEDGHCFTILSKCLCMKRCNAKTLATELLA

```
RESULT 12
LOCUS       HS1043E3      126368 bp    DNA     linear   PRI 04-MAR-2003
DEFINITION  Human DNA sequence from clone RP5-1043E3 on chromosome 6p21.1-21.2,
            complete sequence.
ACCESSION   AL035464
VERSION     AL035464.20  GI:5360970
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 126368)
AUTHORS     Patel R.
TITLE       Direct Submission
JOURNAL     Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 6, 1999 this sequence version replaced gi:5262286.
```

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
RP5-1043E3 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

```
FEATURES             Location/Qualifiers
     source          1..126368
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704E031043"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p21.1-21.2"
                     /clone="RP5-1043E3"
                     /clone_lib="RPCI-5"
BASE COUNT       36745 a   24490 c   25735 g   39498 t
ORIGIN

Query Match           9.1%;  Score 49.2;  DB 9;  Length 126368;
Best Local Similarity 54.5%;  Pred. No. 0.87;
Matches 121; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

Qy     311 ATAACTAATTACAGAGATTAGAGAGAAAATTAAGAGATGAGCAGTATCACACATAAAAAAGTTCTA 370
Db   27667 AGAAATCACTTGTGTGATTTCACCAAAGATAACAAGTATTATTACTGGGTTGTAAATCCTTCAA 27726

Qy     371 CCTTTCTTAAAAGTGTAGCTAAATGTTGTGTTTTTAAATTGGCTTTTAAATTGGCTTTTAATTAC 430
Db   27727 AATTTTTAATGAATACATATAAAATATATTTAAATAAAAAATGGACTCATCCTTTATTG 27786

Qy     431 ACTTTAAATAAGTGTGGCACTTCAAATCCTTTGTTGCAATCCTTGCACTAAGTTTATTTGTGT 490
Db   27787 CTGTTTTTTAACTTGTGTTTTTCATTTAATTGTGACGTCTTTCTTATATCTGTATATAGGTGT 27846

Qy     491 ACTTTTAAATGAAAATGACCTTCTATGGGTCTTTGGTTAAAAAA 532
Db   27847 ACAT--CTCCAACACGAGGTTCTATGGGCCTTTCTATACAAAA 27885
```

```
RESULT 13
LOCUS       AC134847      172706 bp    DNA     linear   HTG 01-OCT-2002
DEFINITION  Mus musculus chromosome UNK clone RP24-171M1, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC134847
```

```
Query Match           17.7%;  Score 95.6;  DB 8;  Length 548;
Best Local Similarity 56.9%;  Pred. No. 7.2e-10;
Matches 312; Conservative 0; Mismatches 174; Indels 62; Gaps 5;

mat_peptide    115..303  /gene="thioni"  /product="gamma-thionin 1"
polyA_signal   366..371  /gene="thioni"
ORIGIN                   /gene="thioni"
BASE COUNT      186 a     76 c    107 g    179 t

Qy     1 ATGGCTCGTCCTCTTGTGCTTCATGGCCATTTGTATCTTGGCAGGATGCTCTCTTTGTTGCC 60
Db    52 ATGGCTGCGTTCCATTTACTTCATCGGCCATTTGCGTCTTGGCAATGACAGTCTTTTTGCT 111

Qy    61 TATGAGGTGCAAGCTAGAGAATGCAAAAACAGAAGCAAAGCAAACATTTCCTTGGAAATATGCATT 120
Db   112 TATGGGTGCAAGGCAAGGAAAATTTGCTGTAACAAAAACCGTTAAAATGTTCT 171

Qy   121 ACCAAACCACCACATGCAGAAAAGCTTGTAT---CAGTGAGAAATTACTACGATGGTCATTGT 177
Db   172 AGTGACCCTCTATGTCAAAAACTCGTATGGAGAGAAAAATATGAGAATGGTCATTGT 231

Qy   178 AGCAAAAATCCTCAGAAGGTGCCTATGTACTAAGCCCATGTGTGTTGAATGAGAGATGACT 237
Db   232 TTCACAATCCTAAGCAAGTGCTTATGCAATGCATGAAGATGTA------- 271

Qy   238 AAAACAGGAGCTGAAAATTTTGGCTGAGGAAGCAAAAACTTTGGCTGCCAGCTTTGCTTGAA 297
Db   272 -----------ATGCTAAAAACTTCTCGCAACACAGAATTGCTTGCT 303

Qy   298 GAAGAGAGATAAATGGATAAACTAAATTAGAGAGAAATTAAGAGATGCAGTATCACACAT 357
Db   304 TGA---AGAGAGTTCCTCAAGCAGTAATTAAGTTTGATTATGAGAGTGCACACA 360

Qy   358 AATAAAGT---TTCTACCTTTCTTAAAAAGTGTAGCTAAAGTTGTGTTTTTAAATTGGCTTT 413
Db   361 AATTAAATAAAGTGTTTGCCTTTCTTAAAGGGTAACTTATTTATAATGTTCTTGGAGTA 420

Qy   414 TAGTAGCGCTTTTATTACACTTAAATAAGTAAGTGTGGCCACTTCAAATCCTTTGTTGCAATCTTGC 473
Db   421 TAGTAGCCATTTGACACATTAAATTAAATTAAAGTTGTGACAACATCCTTCATGTATTCTTTCTC 480

Qy   474 ACTAAGTTTATTTGTGTACTTTAATGAAAATGACCTTCTATGGGTCTTTGGTTAAAAAAA 533
Db   481 ATTAAG---TTTGTTGTTTTAAATGAAAGAGAGATCGGTTTACGGTCTCTATATGTAAAAAA 536

Qy   534 AAAAAAAA 541
Db   537 AAAAAAAAA 544
```

```
VERSION     AC134847.1  GI:23396373
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus
  ORGANISM  Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 172706)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172706)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     --------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ----------- Project Information -----------
            Center project name: M_BB0171M01
            --------------- Summary Statistics ---------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 169864 bases at least Q40
            Consensus quality: 170438 bases at least Q30
            Consensus quality: 170769 bases at least Q20
            Insert size: 164000; agarose-fp
            Insert size: 172784; sum-of-contigs
            Quality coverage: 13.54 in Q20 bases; agarose-fp
            Quality coverage: 10.47 in Q20 bases; sum-of-contigs
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *      1      26163: contig of 26163 bp in length
            *  26164      26263: gap of unknown length
            *  26264      48963: contig of 22700 bp in length
            *  48964      49063: gap of unknown length
            *  49064      82663: contig of 33600 bp in length
            *  82664      82763: gap of unknown length
            *  82764     124741: contig of 41978 bp in length
            * 124742     124841: gap of unknown length
            * 124842     171976: contig of 47135 bp in length
            * 171977     172076: gap of unknown length
            * 172077     172203: contig of 127 bp in length
            * 172204     172303: gap of unknown length
            * 172304     172706: contig of 403 bp in length.
FEATURES             Location/Qualifiers
     source          1..172706
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP24-171M1"
     misc_feature    1..26163
                     /note="assembly_name:Contig12"
     misc_feature    26264..48963
                     /note="assembly_name:Contig13"
     misc_feature    49064..82663
                     /note="assembly_name:Contig14"
     misc_feature    82764..124741
```

```
     misc_feature    124842..171976
                     /note="assembly_name:Contig15"
     misc_feature    172077..172203
                     /note="assembly_name:Contig16"
     misc_feature    172304..172706
                     /note="assembly_name:Contig10"
BASE COUNT    53851 a    32815 c    32725 g    52700 t      615 others
ORIGIN

Query Match          9.1%;  Score 49;  DB 2;  Length 172706;
Best Local Similarity  53.4%;  Pred. No. 0.89;
Matches  103;  Conservative  0;  Mismatches  90;  Indels  0;  Gaps  0;

Qy   286 GCTTTGCTTGAAAGAGAGAGGATAATGATAACTAATTATTAGAGATTAGAAGAAATTAAGGGATGC 345
         ||||| |||||| ||| ||||  |   |    |    |  || || |||| | ||  | ||| ||
Db 23940 GTTTTTGGAGGAGAAATCAGATAGGGGACAACACATTGAAATGAAATAAATAAATTAATATTT 23881
Qy   346 AGTATCACACATAATAAAAGTTTCTRACCTTTCTTAAAAGTCTAGCTAAATGTTGTGTTTTAA 405
         |||||| || ||||| |||||| | ||  |  ||||||| || || ||||| || | || ||
Db 23880 AATAAAAATAAAAATAAATAATAAAATTTACACAATTATTTTTTCTTTTTTGCT 23821
Qy   406 TTGGGCTTTAGTAGCCTTTTATTCACACTTTAAATAAGTGTGGGCACTTCAATCCTTTGTGC 465
         ||||  | ||||| ||| |  |   ||  | |  ||  |  || | |  |  || || ||
Db 23820 TTGGGCTTTTATTATTTTTTTTTCATTTTTTTATTTAGGTATTTAACTCATTTACAITTCC 23761
Qy   466 AATCTTGCACTAA 478
         |||| |||
Db 23760 AAATGCTATACTAA 23748

RESULT 14
AC004153/c
LOCUS       AC004153            156060 bp    DNA    linear    HTG 12-AUG-2000
DEFINITION  Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
            PROGRESS ***, 2 unordered pieces.
ACCESSION   AC004153
VERSION     AC004153.7  GI:9777733
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Plasmodium falciparum (malaria parasite P. falciparum)
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 156060)
  AUTHORS   Hyman,R.W., Fung,E.L., Qin,P., Rowley,D., Mao,J., Tamaki,T.,
            Kurdi,O.B., Conway,A.B. and Davis,R.W.
  TITLE     Plasmodium falciparum 3D7 chromosome 12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 156060)
  AUTHORS   Hyman,R.W., Qin,P., Fung,E.L., Conway,A.B. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-1998) Stanford DNA Sequencing and Technology
            Center, Stanford University, 855 California Avenue, Palo Alto, CA
            94304, USA
COMMENT     On Aug 12, 2000 this sequence version replaced gi:8810454.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *      1     102169: contig of 102169 bp in length
            * 102170     102369: gap of unknown length
            * 102370     156060: contig of 53691 bp in length.
FEATURES             Location/Qualifiers
     source          1..156060
                     /organism="Plasmodium falciparum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5833"
                     /chromosome="12"
                     /clone="PFYAC812"
                     /clone="3D7"
```

BASE COUNT   62615 a   14889 c   15137 g   63219 t   200 others
ORIGIN

Qy   307 ATGGATAACTAATTAGGAGATTAGAAGAAATTAAGGATGCGAGTATCACACATAATAAAGTT 366
            ||| ||||| |||||| |||| |||||| |||||||||||| |||| ||||||||||||||
Db   66714 ATATATATTATAATTAAATAATTAAAAATAATATCATTATATTAAAAGTCGCATTATAAAAATTT 66655

Qy   367 TCTACCTTTCTTAAAAGTGTAGCTAAAATGTTGTGTTTTAAATTGTTTTAGTAGCCTTTTA 426
            |||||| ||| | |||||||||||| ||||| |||||| |||| |||||||||
Db   66654 GTGTAACCCTTTACATTTTTTAATTCATTTTTTTTAAATTGCAATTATTTTAAAAAAAAA 66595

Qy   427 TTACACTTTAAATAAGTGTGGCACTTCAATCCTTTGTGCAATCTTGCACTAAGTTTATTT 486
            ||||| | |||||||| |||||||||||| || |||| ||||||| |||||||||
Db   66594 TTTTA-TTTAATAACTTCTAAACATCTTTTTAAGGTAAAAATATATGCTTAAATTTAATTT 66536

Qy   487 GTGTACTTTAATGAAAATGACCTTCTATGGTGTCTTTGGGTTTAAAAAAAAAAAA 541
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   66535 TTTTTAAATTAAAATACAAGATATATATATTATTATTATTATATAAATAAATAAAA 66481

RESULT 15
LOCUS   AE014848/c
DEFINITION   AE014848   250707 bp   DNA   linear   INV 11-FEB-2003
               Plasmodium falciparum 3D7 chromosome 12, section 5 of 9 of the
               complete sequence.
ACCESSION   AE014848 AE014188
VERSION   AE014848.1 GI:23496770
KEYWORDS   .
SOURCE   Plasmodium falciparum 3D7
  ORGANISM   Plasmodium falciparum 3D7
               Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1 (bases 1 to 250707)
  AUTHORS   Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
               Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
               James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
               Kyes,S., Chan,M.S., Nene,V., Shallom,S.J., Suh,B., Peterson,J.,
               Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
               Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
               Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
               Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
               Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
               Barrell,B.
  TITLE   Genome sequence of the human malaria parasite Plasmodium falciparum
  JOURNAL   Nature 419 (6906), 498-511 (2002)
   PUBMED   12368864
REFERENCE   2 (bases 1 to 250707)
  AUTHORS   Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
               Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
               University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE   3 (bases 1 to 250707)
  AUTHORS   Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
               Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
  TITLE   Direct Submission
  JOURNAL   Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
               University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES   Location/Qualifiers
     source   1..250707
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="genomic DNA"
                     /isolate="3D7"
                     /db_xref="taxon:36329"
                     /chromosome="12"
     repeat_region   108..136
                     /rpt_type=tandem
     repeat_region   372..405
                     /rpt_type=tandem

     repeat_region   375..405
                     /rpt_type=tandem
     repeat_region   477..541
                     /rpt_type=tandem
     repeat_region   671..726
                     /rpt_type=tandem
     repeat_region   1146..1171
                     /rpt_type=tandem
     repeat_region   1522..1551
                     /rpt_type=tandem
     repeat_region   1547..1578
                     /rpt_type=tandem
     repeat_region   1600..1655
                     /rpt_type=tandem
     gene   <1656..>2618
                     /locus_tag="PFL1195w"
     mRNA   <1656..>2618
                     /locus_tag="PFL1195w"
     CDS   1656..2618
                     /locus_tag="PFL1195w"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAN36325.1"
                     /db_xref="GI:23496771"
                     /translation="MLNNLPCNKKYWNSILSAQKYHFCNKWLSKRSNIKKSDICKYF
                     GVSGSYGYSIFSKEGNNNSNVGKNNLHGYEBKCLINNFIKKSNFYFTKWPGKVSKLGG
                     NYFFNKNYNMGFYHMIEEBKGIKKQDIALISSCYERKSIQNKTIENKMKDLTCGYSVQI
                     VLSGKGVKCYEBDINYTPILENYKRNNIKQYAFKNTLKDLKVENRSDHNNNVENTQN
                     VKTISNNTQNNYRTKKVKKLFVRLGITGIVDIDITRVLTWHKKYVNINVDKTGTIVNV
                     YGYNKKKVGVSAYRLYKIVKMVYTMKGGYYYLHHKPKQKIPKKK"
     gene   3159..3191
                     /rpt_type=tandem
     repeat_region   2831..2865
                     /rpt_type=tandem
     repeat_region   3159..3191
                     /rpt_type=tandem
     gene   complement(<3199..>3709)
                     /locus_tag="PFL1200c"
     mRNA   complement(join(<3199..3463,3654..>3709))
                     /locus_tag="PFL1200c"
     CDS   complement(join(3199..3463,3654..3709))
                     /locus_tag="PFL1200c"
                     /note="IMMPfam hit to PF00076, RNA recognition motif.
                     (a.k.a. RRM, RBD, or RNP domain), score 1.4e-13"
                     /codon_start=1
                     /product="hypothetical protein, conserved"
                     /protein_id="AAN36326.1"
                     /db_xref="GI:23496772"
                     /translation="MSRNIRLPAEVSRILYVRNLPYKISADELYDIFGKKYGYGTVR
                     KGNAEGTKGTSFVVYDDIYDAQNALDHLSGFNVAGRYLVVLIYDPVKAQKKEIQEKL
                     KNEK"
     repeat_region   4161..4195
                     /rpt_type=tandem
     repeat_region   4203..4244
                     /rpt_type=tandem
     repeat_region   4547..4584
                     /rpt_type=tandem
     gene   complement(<4677..>102215)
                     /locus_tag="PFL1205c"
     mRNA   complement(join(<4677..4677,4925..4925,6566..6762..9225,
                     9280..>102215))
                     /locus_tag="PFL1205c"
     CDS   complement(join(4677..4677,4925..4925,6566..6762..9225,
                     9280..10215))
                     /locus_tag="PFL1205c"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAN36327.1"
                     /db_xref="GI:23496773"
                     /translation="MLNNNAKKIQLLKKILYRENIKRKQKNNLRVVFNDYINNERIIF
                     LYDKNIFIQIIEFNNKQKSLYIHSDVFKNNKGNGNDQLIIKKGNLKSFVCHKENKHLG
                     EKRKNIKVNIKNVYPHNNKNIYNDANMYNHNQHDISIDTSYKNHHTHQCVKINNYIS
                     NEEIIDKKINQPNENHTRHNDPCDNMKPSTYCLEEKRNYLNVANPVIIQKKRNLEIVKK
                     YSDKWEKHENKQKLIIKYEIKNTPDIINILNKCKEYANYINNNLYDIIYNIINTRDN
                     YLSLKSLILISNSPIKLNYFNETFYNYIFIHILKLCKLNMYNYIFLYXMNFMFLHRLY"

```
SLSIYLQIYIILSCYLKIYKIYVDLKKINEHSSNNYTIHNKEKIYLGDNNEHVLKKYNK
INSQQYHIFYKINFSQIVRQNENNNNVKINNESIDLVFFIIESLTDKIKININLDTLID
ISDPLNDLFFSYFKKIIGSLFLKKVHSEYITFNKEQRKILYKRDSIIQMYKKEQLKK
NDDMSRGGVLKGGVSKGGVFFKQQNDCNINITNKQGHIKNSQKILLKIISKMKRKIRA
RIFIPFIKIYEBKDMKIINKSKDIIKEDTNLLYTFPCNTYKNMIITFDEKSLLIT
LSCFSKGPLLFNDVIYMLTYLMEKKIIKYYNHEEIIIDLMNSYSNMKNEKISNNVFNFL
ARFLFNNKIIHLNDTHIHILNLIIKKCFINQQDVIKYISNFIMYHKPYYKNLKAS
LYLFFHYHFNSLNEQPLANFIFCTLNHKDNNTLYHCNQVFSSITTIIGSNQLKYCKQGVTN
ILDTSILNLKEIYFNKHVDISILNDMKNIVDFVFYFNKSITGNVLKAEYRKIKVSNDK
KKKKKKKKNHIIQVKEKIYNNDNIYDNIYDNIYDNIYDNDNCDNNYDN
VNMSIYTPIAPKCCNKKKPVFFSNMSFYCKIYLMKFSFNIRMFIILKQIVEYLLK
EKYYFKNNNQFIILMVKLLSNVSEIVKLINSLNNNYTYLIYINNKSYVKNKIINNYFL
RILNYINNNILLFLCKANERIILFLSNFVMLKNKYTFEHINYHKIIHTYLEKQIDKVRN
KDTDISSILIFDFLSSYQIDEIMILARCIKMKGQIESSHMHEKRDLEEIHHTMLKNF
LYSQKLRGRBEBKMKYEVTSIMYNEKKKKKKSSIIQNIMKNNFFVTLQLRFEYISKIFKN
KYEKLYMSKVLFNNIEKRMIKRIKIENMDAKILKLNNSPRHINKYNKRYVWNKFCKKWKKR
KKYILIKCTHISQVIDTNNLIDLLIYTLTTIRYYMNNICSQNNRIKNISSFIIYEKF
IIFILNKMVLLKRPQIVANINYYLYYKMNLCIHTLKYVHVRLLYKSYYLLLMYMKFIILR
ILQRNKTFKNYMDFIHFNNNSRTSNNA1QNIEENEIHEHNSTVHNLYNTYPSISQTSYK
YKLQNDVNKNILQNNLHYXNTQFYKNLVIAHHNLYQQFSYLQNASHQLMYFIFYPKLK
GRNMSLIFNGKYIKKEIKRRKKGRINRNMIINTLKLNKTSFLINDAANHKETKKKNTS
ILLKNNIELILTVKKLYILLNHLIQVHVKNKSKSTNIINMNTNNENILIYFFIYSSNNQ
HVPNHEHSSKVGIQKYNNNHTKEVLRKYQKSDKSNYALKQKEGTIIYLKRDTFLNLY
IIQHELHKKLTSPFRLIPVCSGELNKIKNRKNLYVYLTKLFV"
```

repeat_region   4804..4840   /rpt_type=tandem
repeat_region   6659..6685   /rpt_type=tandem
repeat_region   6683..6711   /rpt_type=tandem
repeat_region   7417..7533   /rpt_type=tandem
repeat_region   7417..7532   /rpt_type=tandem
repeat_region   7562..7596   /rpt_type=tandem
repeat_region   8579..8620   /rpt_type=tandem
repeat_region   10342..10382   /rpt_type=tandem
repeat_region   10414..10454   /rpt_type=tandem
repeat_region   10454..10489   /rpt_type=tandem
repeat_region   <10873..>16375   /rpt_type=tandem

gene   join(<10873..11147,11297..12332,12521..16288..>16375)

mRNA   join(10873..11147,11297..12332,12521..16212,16288..>16375)

CDS   join(10873..11147,11297..12332,12521..16212,16288..16375)
   /locus_tag="PFL1210w"
   /codon_start=1
   /note="HMMPfam hit to PF00133, tRNA synthetases class I
   (I, L, M and V), score 2.8e-08"
   /product="hypothetical protein"
   /protein_id="AAN36328.1"
   /db_xref="GI:23496774"
   /translation="MFKYGNFIVIFLIFLSLIIELYLSLHIQKKHIPLDNXNNSFLQ
FINSEHYKYKHNNKYQLNQHKQNIKDVIKESINLPLQFLPTTYNSLDRQKTQQFWNC
KEIYEKRNENIKKHIYANGNNNTKQSYIQHHNTQSCNEEHNINKKKKKKKEYKSDVM
ETKNSHLIKELIKELIEKKKLISKIDDNNNNNNRKHFKYIKTKDYINCHNGNHLINKFEN
IIYEKIKNEKYLKKNKEILKLYKNKILNKIIHDGPPYAANDIHIGHILNKVIKDI
YLXFFLFQNYFVLLIHGPDTHGLPIEINVMKLLKINIQDLALNNSSYIHNQNESLKNE
HVFFKSYINLNKLKNQHKKKQISFLSNIFEKNNIISKDTIEQQKISTFKNLCKSYA
SYFNEQFMSLVISYGIWGYWNYTITYKFYEQIQNKYFRDLLKNKIIYMSNRPIYHS
YATKTVLSSEIIYKRVCNSFYFYFYFFKISDTFLIHLLKEFKENKAINEILEFNQD
DVTFLENYSLVINEIIKNNKENNEGRNIINGNDLFNETNSNDSYLEGQHIISNYILKI
KNIIEKIIKKVKIVLTLTTQMHTIFNNKCLLIHERYLYRIIHHLKYEDEKENQFFICD
KSYNSFSDNLTKYYSKSSPIKEIKNICFQGNSFMSCTYKNFTNNEENNFFIVSDNEI
NDMKAEVSKKIFNENYFIKNNSLSLDVINENVMDENDLDKEYTELVLYKNLEKNIHFI
KSFCTGIVHVAPSHGFTDYNFYYKNNLKKKIYLDISNYQEIEEKKQTRDNNKININIC
KKYETPSNAQSLLNSLDIMKKKTKKLNIEKDIHSLFFYSFYENILFYPYEHSYTYD
WRSHTTVQIKSLLQIYVDIDKICNNIPFYQSIKNIHFINKHVKNNLIKITKDRNEWCI
```

Query Match   9.0%;   Score 48.6; DB 3; Length 250707;
Best Local Similarity 53.2%; Pred. No. 0.99;
Matches 125; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

SRQKYWGVNIFLKDIELDQNQKIIFNNQIMDVWFDSSVSYIVLYMCKHILFHTYFNK
IWKSVKNYNNTYQKGKGKYKNKINYANKNENKLSFNIYDVVDQIMHDEYNIRSMENK
DNILLKNILQKKKIFPDYSKMYKWINKQIQKGQQKKYYSFQWNSIKDIVFSKNNKEFF
KKHFNIFLCCEGIDQIRGWFQSFFFVFFCLNWWINQRKKTRSQLLRNTDKIIDKKGVVI
IKREQIKNNIKHYTLNNNNNDNNNNNNSIFSDDSLCFNNTSSIETNSKRLPSHSYL
PIKNVIVHNYVVDSNNIKMSKSLNNVISPRELFFEKEKDDIPTTSRKDTQKPDDLNK
NTDINNKGNINNNVNNNNDILMKTGKQIDLLKKKNKNKDLNKRFNADIVRLWVCCYNF

Qy 307 ATGGATAACTAATTAGAGGATTAGAAGAAATTAAGGAATGCCAGTATCACACATAATAAAGTT 366
     ||||| |||||| || ||| ||| || | |||| || ||||| || || |||| ||| |||
Db 85617 ATATATATATAAAATATAAAACTATATCAATATATAAAGTCGGCACATTATAAAAAATT 85558

Qy 367 TCTACCTTTCTTAAAAGTGTAGCTAAGTTGGTTGTGTTTAATTGGCTTTTAGTAGCCTTTTA 426
Db 85557 GTGTAAACCCTTTACATTTTAATTCATTTTTTTTTTTAATTGCAATTATTTTAAAAAA 85498

Qy 427 TTACACTTTAAAATAAGTGTGGCACTTCAATCCTTTGTCGACTCTTGCGACTAAGTTATTT 486
Db 85497 TTTTA-TTTTAATAACTTCTAAACATCTTTTAAGGTAAAATATATGCTTAATTTAATTT 85439

Qy 487 GTGTACTTTTAATGAAAAATGACCTTCTATGGTCTTTTGGTTATAAAAAAAAAAAAAGTT 541
Db 85438 TTTTTAAATTAAATACAAGATATATATTTATTATTATTATTAATAATAAAATAAA 85384

Search completed: January 28, 2004, 09:18:02
Job time : 1970 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 28, 2004, 08:46:51 ; Search time 14 Seconds
                  (without alignments)
                  352.700 Million cell updates/sec

Title: US-10-072-809B-18
Perfect score: 552
Sequence: 1 MARSLCFMAPAILLARMLFVA..........LAEEAKTLAAALIEEEIMDN 105

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Database :   SwissProt_41.*

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Total number of hits satisfying chosen parameters: 127863

Searched: 127863 seqs, 47026705 residues

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	529	95.8	105	1	THGF_TOBAC	P32026 nicotiana t
2	468.5	84.9	106	1	THG1_NICPA	O24115 nicotiana p
3	144	26.1	78	1	THG_PETIN	O40901 petunia int
4	143.5	26.0	73	1	THG4_ARATH	Q9ZU8 arabidopsis
5	138.5	25.1	74	1	P332_SOLTU	P20346 solanum tub
6	137.5	24.9	77	1	THG1_ARATH	O39182 arabidopsis
7	137.5	24.9	77	1	THG3_ARATH	Q9zu17 arabidopsis
8	137	24.8	78	1	THGF_HELAN	P82659 helianthus
9	135	24.5	74	1	DEF2_CAPAN	O65740 capsicum an
10	132	23.9	52	1	DEF2_SPIOL	P81571 spinacia ol
11	127.5	23.1	77	1	THG2_ARATH	Q41914 arabidopsis
12	123	22.3	75	1	DEF1_CAPAN	O43413 capsicum an
13	112	20.3	80	1	AFP2_ARATH	O80995 arabidopsis
14	111	20.1	74	1	SRP_SOYBN	Q07502 glycine max
15	110	19.9	80	1	AFP1_BRANA	P30224 brassica na
16	108	19.6	80	1	AFP4_RAPSA	O24331 raphanus sa
17	107	19.4	47	1	THG1_MAIZE	P81008 zea mays (m
18	106.5	19.3	79	1	AFP3_RAPSA	O24332 raphanus sa
19	105	19.0	80	1	AFP3_ARATH	O80994 arabidopsis
20	103	18.7	47	1	SA1_SORBI	Q09198 sorghum bic
21	103	18.7	47	1	SIA3_SORBI	P21925 sorghum bic
22	103	18.7	80	1	AFP1_ARATH	P30224 arabidopsis
23	103	18.7	80	1	AFP2_RAPSA	P30230 raphanus sa
24	102.5	18.6	79	1	AFP3_BRANA	Q39313 brassica na
25	102	18.5	47	1	THCC_VIGUN	P83399 vigna ungui
26	101	18.3	47	1	FAB1_VICFA	P81456 vicia faba
27	101	18.3	47	1	FAB1_VICFA	P81457 vicia faba
28	99	17.9	47	1	THG_HORVU	P20230 hordeum vul
29	99	17.9	48	1	SIA2_SORBI	P21924 sorghum bic
30	97	17.6	80	1	AFP4_ARATH	Q9fi23 arabidopsis
31	94.5	17.1	74	1	DR9_PEA	Q01784 pisum sativ
32	94	17.0	47	1	THG2_WHEAT	P20159 triticum ae
33	92.5	16.8	75	1	1OXD_VIGUN	P18646 vigna ungui
34	88.5	16.0	51	1	AF2A_SINAL	P30232 sinapis alb
35	88	15.9	47	1	THG1_WHEAT	P20158 triticum ae
36	86.5	15.7	72	1	D230_PEA	Q01783 pisum sativ
37	85.5	15.5	46	1	AX2_BETVU	P82010 beta vulgar
38	85.5	15.5	51	1	AFP1_SINAL	P30231 sinapis alb
39	84.5	15.3	46	1	AX1_BETVU	P81493 beta vulgar
40	78	14.1	47	1	SIA1_SORBI	P21923 sorghum bic
41	77	13.9	70	1	DMYC_DROME	P41964 drosophila
42	76	13.8	47	1	PSD2_PEA	P81930 pisum sativ
43	76	13.8	47	1	THI22_MAIZE	P81009 zea mays (m
44	73.5	13.3	297	1	SGS4_DROME	O00725 drosophila
45	68	12.3	322	1	FSA_BRARE	Q9yhv4 brachydanio

ALIGNMENTS

RESULT 1
ID   THGF_TOBAC   STANDARD;   PRT;   105 AA.
AC   P32026;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flower-specific gamma-thionin precursor.
GN   FST.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. W38; TISSUE=Flower buds;
RX   MEDLINE=92357021; PubMed=1495489;
RA   Gu Q., Kawata E.E., Morse M.J., Wu H.-M., Cheung A.Y.;
RT   "A flower-specific cDNA encoding a novel thionin in tobacco.";
RL   Mol. Gen. Genet. 234:89-96(1992).
CC   -!- FUNCTION: INVOLVED IN FLORAL ORGANOGENESIS. MAY PLAY A PROTECTIVE
CC   ROLE IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM
CC   POTENTIAL PATHOGEN ATTACK.
CC   -!- SUBCELLULAR LOCATION: POSSIBLY THE CELL WALL OR VACUOLE.
CC   -!- TISSUE SPECIFICITY: FLOWER. FOUND IN PETALS, STAMEN AND PISTILS,
CC   BUT NOT IN SEPALS. IN PARTICULAR, ACCUMULATION IN A CONFIGURATION
CC   SURROUNDING THE INNER REPRODUCTIVE WHORLS.
CC   -!- DEVELOPMENTAL STAGE: ACCUMULATES IN DEVELOPING FLOWERS AND ITS
CC   LEVEL DROPS AS FLOWERS MATURE.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Z11748; CAA77806.1; -.
DR   PIR; S23574; S23574.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Puctthionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Plant toxin; Signal.
FT   SIGNAL   1   25   POTENTIAL.
FT   CHAIN   26   105   FLOWER-SPECIFIC GAMMA-THIONIN.
FT   DISULFID   28   72   BY SIMILARITY.
FT   DISULFID   39   59   BY SIMILARITY.
FT   DISULFID   45   66   BY SIMILARITY.
FT   DISULFID   49   68   BY SIMILARITY.
SQ   SEQUENCE   105 AA; 11750 MW; DA7F451F065LBC3 CRC64;

```
Qy   Query Match
     Best Local Similarity   95.8%;   Score 529;   DB 1;   Length 105;
     Matches 100;   Conservative    1;   Mismatches    4;   Indels    0;   Gaps    0;

Qy     1  MARSLCFMAFAILARMLFVAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db     1  MARSLCFMAFAILAWMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60

Qy    61  KILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIMDN 105
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    61  KILRRCLCTKPCVFDEKMIKTGAEILVEEAKTLAAALLEEEIMDN 105
```

```
RESULT 2
ID   THG1_NICPA              STANDARD;      PRT;   106 AA.
THG1_NICPA
AC   O24115;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-thionin 1 precursor.
GN   THIO1.
OS   Nicotiana paniculata.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=62141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Komori T., Yamada S., Imaseki H.;
RT   "A cDNA clone for gamma-thionin from Nicotiana paniculata.";
RL   (In) Plant Gene Register PGR97-132.
CC   -!- INDUCTION: By salt stress.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AB005250; BAA21325.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Plant toxin; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    106       GAMMA-THIONIN 1.
FT   DISULFID     29     73       BY SIMILARITY.
FT   DISULFID     40     60       BY SIMILARITY.
FT   DISULFID     46     67       BY SIMILARITY.
FT   DISULFID     50     69       BY SIMILARITY.
SQ   SEQUENCE   106 AA;   11722 MW;   26985OE116D775A4 CRC64;
```

```
Qy   Query Match
     Best Local Similarity   84.9%;   Score 468.5;   DB 1;   Length 106;
     Matches   89;   Conservative    8;   Mismatches    8;   Indels    1;   Gaps    1;

Qy     1  MARSLCFMAFAVLAMMLFVAYEVQOARE-CKTESNTFPGICITKPPCRKACISEKFTDGHC 59
          |||||||||||:|:||||||||||| :||||||||||||||||:|||||||||||||||
Db     1  MARSLCFMAFAVLANMLFVAYEVQAKSTCKAESNTFPGICITKPPCRKACISEKFTDGHC 60

Qy    60  SKILRRCLCTKPCVFDEKMTKTGAEILAEEAKTLAAALLEEEIMDN 105
          :|||:|:||||||||||||||||||||||||||||||||||||||
Db    61  SKILRRCICKXPCVFDGKRMIQTGAENIAEEAETLAAALLEEEMMDN 106
```

```
RESULT 3
ID   THG_PETIN               STANDARD;      PRT;   78 AA.
THG_PETIN
AC   Q40501;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-thionin homolog PPT precursor.
OS   Petunia integrifolia (violet-flowered petunia) (Petunia inflata).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pistil;
RX   MEDLINE=95036017; PubMed=7948892;
RA   Karunanandaa B., Singh A., Kao T.H.;
RT   "Characterization of a predominantly pistil-expressed gene encoding a
RT   gamma-thionin-like protein of Petunia inflata.";
RL   Plant Mol. Biol. 26:459-464(1994).
CC   -!- FUNCTION: May be involved in the defense of the pistil against
CC       pathogen infection.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE PISTIL DURING
CC       ALL STAGES OF FLOWER DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L71173; AAA64740.1; -.
DR   PIR; S52634; S52634.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32     78       GAMMA-THIONIN HOMOLOG PPT.
FT   DISULFID     34     78       BY SIMILARITY.
FT   DISULFID     45     65       BY SIMILARITY.
FT   DISULFID     51     72       BY SIMILARITY.
FT   DISULFID     55     74       BY SIMILARITY.
SQ   SEQUENCE   78 AA;   8799 MW;   263DB3AC90619265 CRC64;
```

```
Qy   Query Match
     Best Local Similarity   35.9%;   Score 144;   DB 1;   Length 78;
     Matches   28;   Conservative   14;   Mismatches   30;   Indels    6;   Gaps    1;

Qy     1  MARSLCFMAFAILARMLFVAYEV-------QARECKTESNTFPGICITKPPCRKACISEKF 54
          |||||||||: :|| |||:::        |:||||:|||||||||||||||:||||||:
Db     1  MGRSIRLPATFLIAMLFLSTEMGPMTSAEARTCESQSHRFHGTCVRESNCASVCQTEGF 60

Qy    55  TDGHCSKILRRCLCTKPC 72
          |||||:||||:|:|| |
Db    61  IGGNCRAFRRCFCTRNC 78
```

```
RESULT 4
ID   THG4_ARATH              STANDARD;      PRT;   73 AA.
THG4_ARATH
AC   Q9ZUL8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-thionin homolog At2g02140 precursor.
GN   AT2G02140 OR F5O4.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
```

```
DR   EMBL; AC005936; AAC97220.1; -.
DR   PIR; D84433; D84433.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal; Multigene family.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29     73       GAMMA-THIONIN HOMOLOG AT2G02140.
FT   DISULFID     31     73       BY SIMILARITY.
FT   DISULFID     42     62       BY SIMILARITY.
FT   DISULFID     48     67       BY SIMILARITY.
FT   DISULFID     52     69       BY SIMILARITY.
SQ   SEQUENCE   73 AA;   7718 MW;   DA27D295AC37393DA CRC64;

Query Match       26.0%;  Score 143.5; DB 1; Length 73;
Best Local Similarity   40.0%;  Pred. No. 1.2e-08;
Matches  30; Conservative  13; Mismatches  27; Indels   5; Gaps   2;

Qy    1 MARSLCFMAFAILARMLFVAY----EVQARECKTESNTFPGICITKPPCRKACISEKFTDG 57
             :  :  :.: | :|: : |  |  |::||  :||
Db    1 MKLSLRISALLMSVMLLFATGMGPVEARTCESPSNXFQGVCLNSQSCAXACPSEGFSGG 60

Qy   58 HCSKILRRCLCTKPC 72
           ||:||  ||||||
Db   61 RCSSL-RCYCSKAC 73
```

RESULT 5
P322_SOLTU

```
ID   P322_SOLTU    STANDARD;    PRT;   74 AA.
AC   P20346;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable protease inhibitor P322 precursor.
```

```
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Bintje;
RA   Stiekema W.J., Heidekamp F., Dirkse W.G., van Beckum J., de Haan P.,
RA   ten Bosch C., Louwerse J.D.;
RT   "Molecular cloning and analysis of four potato tuber mRNAs.";
RL   Plant Mol. Biol. 11:255-269(1988).
CC   -!- TISSUE SPECIFICITY: TUBER.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
```

```
DR   EMBL; X13180; CAA31577.1; -.
DR   PIR; S05594; S05594.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Serine protease inhibitor; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     74       PROBABLE PROTEASE INHIBITOR P322.
FT   DISULFID     30     74       BY SIMILARITY.
FT   DISULFID     41     61       BY SIMILARITY.
FT   DISULFID     47     68       BY SIMILARITY.
FT   DISULFID     51     70       BY SIMILARITY.
SQ   SEQUENCE   74 AA;   8414 MW;   FAFE1A87038F4B5A CRC64;

Query Match       25.1%;  Score 138.5; DB 1; Length 74;
Best Local Similarity   40.3%;  Pred. No. 4.2e-08;
Matches  29; Conservative  11; Mismatches  25; Indels   7; Gaps   2;

Qy    7 FMAFAILARMLFVAYEV-------QARECKTESNTFPGICITKPPCRKACISEKFTDGHCS 60
             || :.:| |: :|.  ||:||  :||
Db    4 FATFFLIA-MLVVATKMGPMRIAEARHCESLSHRFKGPCTRDSNCASVCETERFSGGNCH 62

Qy   61 KILRRCLCTKPC 72
           |||| ||||||
Db   63 GFRRCFCTKPC 74
```

RESULT 6
THG1_ARATH

```
ID   THG1_ARATH    STANDARD;    PRT;   77 AA.
AC   Q39182; Q42011;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-thionin homolog At2g02100 precursor.
GN   AT2G02100 OR F5O4.13.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. C24; TISSUE=Flower buds;
RA   Yu D.Y., Quigley F., Mache R.;
RT   "Isolation and expression of a cDNA encoding protease inhibitor
```

Left column

RT   precursor.";
RL   Submitted (NOV-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Talion L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RL   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL   thaliana.";
RL   Nature 402:761-768(1999).
RN   [3]
RP   SEQUENCE OF 1-75 FROM N.A.
RC   STRAIN=cv. C24; TISSUE=Flower buds;
RA   Mache R., Quigley F., Thomas F., Yu D.Y.;
RL   Submitted (NOV-1992) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: Secreted (Potential).
-!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; X69919; CAA49892.1; -.
DR   EMBL; AC005936; AAC97223.1; -.
DR   EMBL; Z18455; CAA79189.1; -.
DR   PIR; S30578; S30578.
HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal; Multigene family.
FT   SIGNAL   1   30   POTENTIAL.
FT   CHAIN   31   77   GAMMA-THIONIN HOMOLOG AT2G02100.
FT   DISULFID   33   77   BY SIMILARITY.
FT   DISULFID   44   64   BY SIMILARITY.
FT   DISULFID   50   71   BY SIMILARITY.
FT   DISULFID   54   73   BY SIMILARITY.
SQ   SEQUENCE   77 AA; 8524 MW; EC04E60C3610965 CRC64;

RESULT 7
Query Match   33.8%; Score 137.5; DB 1; Length 77;
Best Local Similarity   24.9%;
Matches   26; Conservative   16; Mismatches   30; Indels   5; Gaps   1;

Qy   1 MARSLCFMAFAILARMLFVA-----YEVQARBCKTESNTFPGICITKPPCRKACISEKFT 55
              :|: :| ||||: ||| : :|:: |||::|:|:| | | |||:||::|:||
Db   1 MKLSMRLISAVLIMFMIFVAIGMGPVTVEARTCESQSHRFKGTCVSASNCANVCHNEGFV 60

Qy   56 DGHCSKILRRCLCTKPC 72
              || |||||::|||:|||
Db   61 GGNCRGFRRCFCTRHC 77

ID   THG3_ARATH STANDARD; PRT; 77 AA.
AC   Q9ZU17;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)

Right column

DE   Gamma-thionin homolog At2g02130 precursor.
GN   At2G02130 OR F5O4.10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Talion L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RL   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL   thaliana.";
RL   Nature 402:761-768(1999).
-!- SUBCELLULAR LOCATION: Secreted (Potential).
-!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; AC005936; AAC97221.1; -.
PIR; CB4433; CB4433.
HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal; Multigene family.
FT   SIGNAL   1   30   POTENTIAL.
FT   CHAIN   31   77   GAMMA-THIONIN HOMOLOG AT2G02130.
FT   DISULFID   33   77   BY SIMILARITY.
FT   DISULFID   44   64   BY SIMILARITY.
FT   DISULFID   50   71   BY SIMILARITY.
FT   DISULFID   54   73   BY SIMILARITY.
SQ   SEQUENCE   77 AA; 8544 MW; 0C80DC561971262E CRC64;

RESULT 8
Query Match   36.4%; Score 137.5; DB 1; Length 77;
Best Local Similarity   24.9%;
Matches   28; Conservative   15; Mismatches   29; Indels   5; Gaps   1;

Qy   1 MARSLCFMAFAILARMLFVA-----YEVQARBCKTESNTFPGICITKPPCRKACISEKFT 55
              :|: :| ||||: ||| : :|:: |||::|:|:| | | |||:||::|:||
Db   1 MKLSVRFISAALLEMVFIATGMGPVTVEARTCESKSHRFKGPCVSTHNCANVCHNEGFG 60

Qy   56 DGHCSKILRRCLCTKPC 72
              || |||||::|||:|||
Db   61 GGKCRGFRRCYCTRHC 77

ID   THGF_HELAN STANDARD; PRT; 78 AA.
AC   P82659;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flower-specific gamma-thionin precursor (Defensin SD2).
GN   SD2.

OS   Helianthus annuus (Common sunflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC   Heliantheae; Helianthus.
OX   NCBI_TaxID=4232;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Flower;
RA   Urdangarin M.C., de la Canal L.;
RT   "A defensin gene expressed in sunflower inflorescence.";
RL   Plant Physiol. Biochem. 38:253-258(2000).
-!- FUNCTION: MAY PLAY A PROTECTIVE ROLE IN FLOWERS BY PROTECTING THE
     REPRODUCTIVE ORGANS FROM POTENTIAL PATHOGEN ATTACK.
-!- SUBCELLULAR LOCATION: CELL WALL OR VACUOLE (POTENTIAL).
-!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN FLOWERS AND TO A LESSER
     EXTENT IN LEAVES. LOWER LEVELS IN HYPOCOTYLS. NO EXPRESSION IN
     ROOTS AND COTYLEDONS.
-!- DEVELOPMENTAL STAGE: EXPRESSED PROGRESSIVELY DURING FLOWER
     DEVELOPMENT REACHING THE HIGHEST LEVEL IN THE MATURE FERTILIZED
     FLOWER STAGE.
-!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.

DR   EMBL; AF178634; AAF72042.1; -.
DR   EMBL; AF141131; AAF66591.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal.
FT   SIGNAL   1   20   POTENTIAL.
FT   CHAIN   21   78   FLOWER-SPECIFIC GAMMA-THIONIN.
FT   DISULFID   34   78   BY SIMILARITY.
FT   DISULFID   45   65   BY SIMILARITY.
FT   DISULFID   51   72   BY SIMILARITY.
FT   DISULFID   55   74   BY SIMILARITY.
SQ   SEQUENCE   78 AA; 8627 MW; 1AD6A9858214 9A2D CRC64;

Query Match   24.8%; Score 137; DB 1; Length 78;
Best Local Similarity   35.9%; Pred. No. 6.4e-08;
Matches   28; Conservative   12; Mismatches   32; Indels   6; Gaps   1;

Qy   1 MARSLCFMAFAILARMLFVAYE-------VQARECKTESNTPFGICITKPPCRKACISEKF 54
                                              ||| ||| ||| ||
Db   1 MKSSMKMFAALLLVVMCLLANEMGGPLVVEARTCESQSHKFKGTCLSDTNCANVCHSERF 60

Qy   55 TDGHCSKILRRCLCTKPC 72
                       ||| ||
Db   61 SGGKCRGFRRRCFCTTHC 78

RESULT 9
ID   DEF2_CAPAN   STANDARD;   PRT;   74 AA.
AC   O65740;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Defensin J1-2 precursor.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=cv. Yolo Wonder; TISSUE=Fruit;
RX   MEDLINE=97037730; PubMed=8883377;
RA   Meyer B., Houlne G., Pozueta-Romero J., Schantz M.L., Schantz R.;
RT   "Fruit-specific expression of a defensin-type gene family in bell
RT   pepper. Upregulation during ripening and upon wounding.";
RL   Plant Physiol. 112:615-622(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Yolo Wonder; TISSUE=Fruit;
RX   MEDLINE=99005242; PubMed=9790581;
RA   Houlne G., Meyer B., Schantz R.;
RT   "Alteration of the expression of a plant defensin gene by exon
RT   shuffling in bell pepper (Capsicum annuum L.).";
RL   Mol. Gen. Genet. 259:504-510(1998).
-!- FUNCTION: PLANT DEFENSE PEPTIDE WITH ANTIFUNGAL ACTIVITY AGAINST
     F.OXYSPORUM AND B.CINEREA.
-!- SUBUNIT: Monomer (By similarity).
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
-!- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIPENING.
-!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.

DR   EMBL; X95730; CAA65046.1; -.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Fungicide; Signal.
FT   SIGNAL   1   27   BY SIMILARITY.
FT   CHAIN   28   74   DEFENSIN J1-2.
FT   DISULFID   30   74   BY SIMILARITY.
FT   DISULFID   41   61   BY SIMILARITY.
FT   DISULFID   47   68   BY SIMILARITY.
FT   DISULFID   51   70   BY SIMILARITY.
SQ   SEQUENCE   74 AA; 8249 MW; D92D8F06F39E1552 CRC64;

Query Match   24.5%; Score 135; DB 1; Length 74;
Best Local Similarity   34.9%; Pred. No. 9.9e-08;
Matches   22; Conservative   12; Mismatches   29; Indels   0; Gaps   0;

Qy   10 FAILARMLFVAYEVQARECKTESNTPFGICITKPPCRKACISEKFTDGHCSKILRRCLCLT 69
              | | | || ||| ||| ||| |||
Db   12 FLMVVMLVFATGMVAEARTCESQSHRKGLCFSKSNGSVCHTEGFNGGHCRGFRRRCFCT 71

Qy   70 KPC 72
Db   72 RHC 74

RESULT 10
ID   DEF2_SPIOL   STANDARD;   PRT;   52 AA.
AC   P81571;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Antimicrobial peptide D2 (So-D2) (Defensin D2) (Fragment).
OS   Spinacia oleracea (Spinach).

OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   STRAIN=cv. Matador; TISSUE=Leaf;
RX   MEDLINE=9833863; PubMed=9762899;
RA   Segura A., Moreno M., Molina A., Garcia-Olmedo F.;
RT   "Novel defensin subfamily from spinach (Spinacia oleracea).";
RL   FEBS Lett. 435:159-162(1998).
CC   -!- FUNCTION: Antimicrobial peptide. Active against Fusarium spp.,
CC   Gram-positive and Gram-negative bacterial pathogens.
CC   -!- SUBCELLULAR LOCATION: Cell wall.
CC   -!- TISSUE SPECIFICITY: Distributed in the epidermal cell layer of
CC   leaves and in the subepidermal layer region of stems. Not in
CC   roots.
CC   -!- DEVELOPMENTAL STAGE: Present throughout the life of the leaf.
CC   -!- MASS SPECTROMETRY: MW=5804; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY. GROUP IV.
HSSP; P20230; 1GPT.
InterPro; IPR002118; Gamma-thionin.
InterPro; IPR003614; Knot1.
Pfam; PF00304; Gamma-thionin; 1.
ProDom; PD002594; G_Purothionin; 1.
SMART; SM00505; Knot1; 1.
PROSITE; PS00940; GAMMA_THIONIN; 1.
NON_TER   52
NON_TER   52
SQ   SEQUENCE   52 AA; 5804 MW; F2F736E3ADF8FEBB CRC64;

Query Match   23.9%; Score 132; DB 1; Length 52;
Best Local Similarity   47.9%; Pred. No. 1.5e-07;
Matches 23; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

Qy   25 ARECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC 72
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   5 SRKCKIPSKTFKGICTRDSNCDTSCRYEGYPAGDCKGIRRRCMCSKPC 52

RESULT 11
ID   THG2_ARATH   STANDARD;   PRT;   77 AA.
AC   Q41914; Q9ZUL6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-thionin homolog At2g02120 precursor.
GN   At2G02120 OR F5O4.11.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL   Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
EMBL; Z17665; CAA79029.1; -.
EMBL; AC005936; AAC97222.1; -.
PIR; B84433; B84433.
HSSP; P20230; 1GPT.
InterPro; IPR002118; Gamma-thionin.
InterPro; IPR003614; Knot1.
Pfam; PF00304; Gamma-thionin; 1.
ProDom; PD002594; G_Purothionin; 1.
SMART; SM00505; Knot1; 1.
PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Signal; Multigene family.
FT   SIGNAL   1   30   POTENTIAL.
FT   CHAIN   31   77   GAMMA-THIONIN HOMOLOG AT2G02120.
FT   DISULFID   31   77   BY SIMILARITY.
FT   DISULFID   33   64   BY SIMILARITY.
FT   DISULFID   44   71   BY SIMILARITY.
FT   DISULFID   50   73   BY SIMILARITY.
FT   DISULFID   54   73   BY SIMILARITY.
FT   CONFLICT   19   19   V -> D (IN REF. 1).
FT   CONFLICT   69   69   R -> G (IN REF. 1).
SQ   SEQUENCE   77 AA; 8578 MW; FB7BE5AF727A9A45 CRC64;

Query Match   23.1%; Score 127.5; DB 1; Length 77;
Best Local Similarity   33.8%; Pred. No. 6.4e-07;
Matches 26; Conservative 12; Mismatches 34; Indels 5; Gaps 1;

Qy   1 MARSLCFMAFAILARMLFVA-----YEVQARECKTESNTFPGICITKPPCRKACISEKFT 55
         |: ||: | ||||||||::| |||||||||||||||||||
Db   1 MKFSMRLISAVLFLVMIFVATGMGPVTEARTCASQSQRFKGKCVSDINCENVCHNEGFP 60

Qy   56 DGHCSKILRRCLCTKPC 72
         | |||||||||||
Db   61 GGDCRGFRRCFCTRNC 77

RESULT 12
ID   DEF1_CAPAN   STANDARD;   PRT;   75 AA.
AC   Q43413;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Defensin J1-1 precursor.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 28-69, AND CHARACTERIZATION.
RC   STRAIN=cv. Yolo Wonder; TISSUE=Fruit;
RX   MEDLINE=97037730; PubMed=8883377;
RA   Meyer B., Houlne G., Pozueta-Romero J., Schantz M.L., Schantz R.;
RT   "Fruit-specific expression of a defensin-type gene family in bell
RT   pepper. Upregulation during ripening and upon wounding.";
RL   Plant Physiol. 112:615-622(1996).
CC   -!- FUNCTION: PLANT DEFENSE PEPTIDE WITH ANTIFUNGAL ACTIVITY AGAINST
CC   F.OXYSPORUM AND B.CINEREA.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.

```
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ORANGE AND RED RIPE FRUIT AND TO
CC       A LESSER EXTENT IN MATURE, GREEN FRUIT. PRESENT IN TRACE IN YOUNG,
CC       GREEN FRUIT.
CC   -!- DEVELOPMENTAL STAGE: TRANSCRIPTS REMAINS VERY LOW DURING FRUIT
CC       DEVELOPMENT AND DRAMATICALLY INCREASES DURING RIPENING.
CC   -!- INDUCTION: By wounding.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X95363; CAA64653.1; -.
DR   HSSP; P20158; 1GPS.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Fungicide; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28     75       DEFENSIN J1-1.
FT   DISULFID     30     74       BY SIMILARITY.
FT   DISULFID     41     61       BY SIMILARITY.
FT   DISULFID     47     68       BY SIMILARITY.
FT   DISULFID     51     70       BY SIMILARITY.
SQ   SEQUENCE   75 AA;   8118 MW;  6E2A8CE76144A51B CRC64;
```

Query Match   22.3%;   Score 123; DB 1; Length 75;
Best Local Similarity   35.8%;   Pred. No.1.9e-06;
Matches   24; Conservative   11; Mismatches   30; Indels   2; Gaps   1;

```
Qy    8  MAFAILARMLFVAYEV--QAEBCKIESNTFPGICITKPPCRKACISEKFTDGHCSKILRR  65
             ||  |:||   ::| | ||  |   :|| |||| ||||| |  : |||| |||
Db    8  VATIFLMLLVFATDWMAEAXICEALSGNFKGLCLSSRDCGNVCREGFTDGSCIGFRLQ  67
```

Qy   66 CLCTKPC 72
         ||||| |
Db   68 CFCTKPC 74

```
RESULT 13
ID   APP2_ARATH     STANDARD;     PRT;    80 AA.
AC   O80995;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable cysteine-rich antifungal protein At2g26010 precursor (AFP).
GN   AT2G26010 OR T1P18.18.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
```

```
RL   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC
CC       CATIONS (BY SIMILARITY).
CC   -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AC004747; AAC31234.1; -.
DR   PIR; T02622; T02622.
DR   HSSP; P30231; 1AYJ.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Fungicide; Signal; Multigene family;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     29
FT   CHAIN        30     80       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   MOD_RES      30     30       PROBABLE CYSTEINE-RICH ANTIFUNGAL PROTEIN
FT                                AT2G26010.
FT   DISULFID     33     80       BY SIMILARITY.
FT   DISULFID     44     65       BY SIMILARITY.
FT   DISULFID     50     74       BY SIMILARITY.
FT   DISULFID     54     76       BY SIMILARITY.
SQ   SEQUENCE   80 AA;   8580 MW;  99E1E0D4443AD67B CRC64;
```

Query Match   20.3%;   Score 112; DB 1; Length 80;
Best Local Similarity   32.5%;   Pred. No.2.9e-05;
Matches   26; Conservative   16; Mismatches   30; Indels   8; Gaps   4;

```
Qy    1  MARSLCFMAFAILARMLFVAYE-----VQARE-CKTESNTFPGICITKPPCPCRKACIS-EKF  54
            |  ||| || |||| | :| |     |  |   |  |||| ||||| ||||||| |   |||
Db    1  MAKSRAIITFLFRALVLFAAFEAPIMVERAQKLCEKPSGTWSGVGGNSNACKNQCINLEGA  60
```

Qy   55 TDGHCSKIL--RRCLCTKPC 72
         | | || | || ||| |
Db   61 KHGSCNYVFPAHKCICYFPC 80

```
RESULT 14
ID   SRP_SOYBN     STANDARD;     PRT;    74 AA.
AC   Q07502;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   8.4 kDa sulfur-rich protein precursor (SE60 protein) (Probable
DE   proteinase inhibitor P322).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Paldal; TISSUE=Seed;
RX   MEDLINE=94103312; PubMed=8278516;
RA   Choi Y., Choi Y.D., Lee J.S.;
RT   "Nucleotide sequence of a cDNA encoding a low molecular weight
RT   sulfur-rich protein in soybean seeds.";
RL   Plant Physiol. 101:699-700(1993).
CC   -!- FUNCTION: MAY FUNCTION AS A SULFUR-RICH STORAGE PROTEIN OR A
```

Left column

DR   EMBL; Z13956; CAA78359.1; -.
DR   PIR; S24965; S24965.
DR   HSSP; P20230; 1GPT.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Signal.
FT   SIGNAL   1   28
FT   CHAIN   29   74   8.4 kDa SULFUR-RICH PROTEIN.
FT   DISULFID   31   74   BY SIMILARITY.
FT   DISULFID   42   62   BY SIMILARITY.
FT   DISULFID   48   68   BY SIMILARITY.
FT   DISULFID   52   70   BY SIMILARITY.
SQ   SEQUENCE   74 AA;   8437 MW;   38E93D064D615CE8 CRC64;

Query Match   20.1%;
Best Local Similarity   34.7%;   Score 111; DB 1; Length 74;
Matches 26; Conservative 14; Mismatches 29; Indels 6; Gaps 3;

Qy   1   MARSLCFMAFAILARMLF---VAYEVQARECKTESNTFPGICITKPPCRKACISEKFTDG 57
Db   3   MRKSCGF--FFLLLLLVFASQVVVQTEGRVCESQSHGFHGLCNRDHNCALVCRNEGFSGG 60

Qy   58   HCSKILRCLCTKPC 72
Db   61   RC-KRSRRCPCTRIC 74

RESULT 15
AFP1_BRANA
ID   AFP1_BRANA   STANDARD;   PRT;   80 AA.
AC   P30225; Q41163;
DT   01-APR-1993 (Rel. 25, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cysteine-rich antifungal protein 1 precursor (AFP1).
GN   AFP1.
OS   Brassica napus (Rape), and
OS   Raphanus sativus (Radish).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708, 3726;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=R.sativus; STRAIN=cv. Ronde Rode Kleine Witpunt; TISSUE=Seed;
RX   MEDLINE=95299950; PubMed=7780308;
RA   Terras F.R.G., Eggermont K., Kovaleva V., Raikhel N.V., Osborn R.W.,
RA   Kester A., Rees S.B., Torrekens S., Van Leuven F., Vanderleyden J.,
RA   Cammue B.P.A., Broekaert W.F.;
RT   "Small cysteine-rich antifungal proteins from radish: their role in
RT   host defense.";
RL   Plant Cell 7:568-573(1995).
RN   [2]
RP   SEQUENCE OF 30-73.
RC   SPECIES=B.napus; TISSUE=Seed;
RX   MEDLINE=93138130; PubMed=8422949;
RA   Terras F.R.G., Torrekens S., Van Leuven F., Osborn R.W.,
RA   Vanderleyden J., Cammue B.P.A., Broekaert W.F.;

Right column

RT   "A new family of basic cysteine-rich plant antifungal proteins from
RT   Brassicaceae species.";
RL   FEBS Lett. 316:233-240(1993).
RN   [3]
RP   SEQUENCE OF 30-73.
RC   SPECIES=R.sativus; TISSUE=Seed;
RX   MEDLINE=92348373; PubMed=1639777;
RA   Terras F.R.G., Schoofs H.M.E., de Bolle M.F.C., van Leuven F.,
RA   Rees S.B., Vanderleyden J., Cammue B.P.A., Broekaert W.F.;
RT   "Analysis of two novel classes of plant antifungal proteins from
RT   radish (Raphanus sativus L.) seeds.";
RL   J. Biol. Chem. 267:15301-15309(1992).
CC -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC
CC   CATIONS.
CC -!- SUBUNIT: FORMS OLIGOMERS IN ITS NATIVE STATE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
CC
DR   EMBL; U18557; AAA69541.1; -.
DR   PIR; T10176; T10176.
DR   HSSP; P30231; 1AYJ.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G_Purothionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Fungicide; Signal; Multigene family;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL   1   29
FT   CHAIN   30   80   CYSTEINE-RICH ANTIFUNGAL PROTEIN 1.
FT   MOD_RES   30   30   PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID   33   80   BY SIMILARITY.
FT   DISULFID   44   65   BY SIMILARITY.
FT   DISULFID   50   70   BY SIMILARITY.
FT   DISULFID   54   76   BY SIMILARITY.
SQ   SEQUENCE   80 AA;   8734 MW;   05B90FAACBDA6C2B CRC64;

Query Match   19.9%;
Best Local Similarity   32.5%;   Score 110; DB 1; Length 80;
Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4;

Qy   1   MARSLCFMAFAILARMLFVAYE-----VQARE-CKTESNTFPGICITKPPCRKACIS-EKF 54
Db   1   MAKFASIIALLFAALVLFAAFEAPTMVEAQKLCERPSGTWSGVCGNNNACKNQCINLEKA 60

Qy   55   TDGHCSKIL--RRCLCTKPC 72
Db   61   RHGSCNYVFPAHKCICYFPC 80

Search completed: January 28, 2004, 08:47:20
Job time : 16 secs